

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C07K 14/47, C12N 5/10, A61K 38/17</b>	<b>A2</b>	(11) International Publication Number: <b>WO 98/25962</b>
		(43) International Publication Date: 18 June 1998 (18.06.98)

(21) International Application Number: PCT/US97/23224

(22) International Filing Date: 12 December 1997 (12.12.97)

(30) Priority Data:

08/766,263 13 December 1996 (13.12.96) US  
08/989,232 11 December 1997 (11.12.97) US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

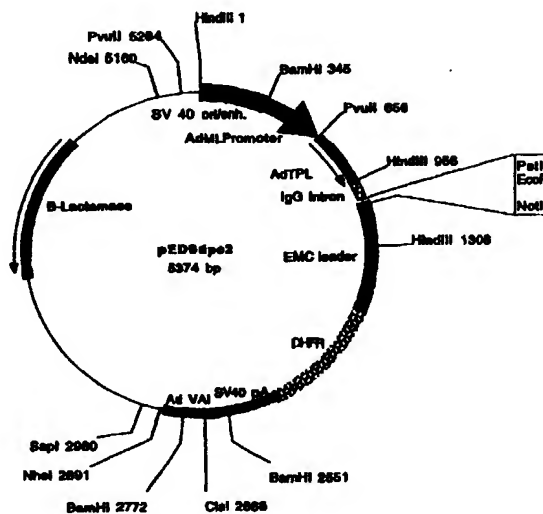
Published

*Without international search report and to be republished upon receipt of that report.*

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED8dpc2  
Plasmid size: 5374 bp

Comments/References: pED8dpc2 is derived from pED8dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), MAR 19: 4485-4490.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	FS	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5

10

## SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

15

This is a continuation-in-part of Ser. No. 60/XXX,XXX, filed December 13, 1996 (converted to provisional application from non-provisional application Ser. No. 08/766,263), which is incorporated by reference herein.

20

### FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

25

### BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

40

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 5 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 22 to nucleotide 462;
- (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AJ1\_1 deposited under accession number 10 ATCC 98278;
- (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AJ1\_1 deposited under accession number ATCC 98278;
- (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AJ1\_1 deposited under accession number ATCC 15 98278;
- (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AJ1\_1 deposited under accession number ATCC 98278;
- (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
- 20 (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
- (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
- (j) a polynucleotide which encodes a species homologue of the protein 25 of (g) or (h) above ; and
- (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 22 to nucleotide 462; the nucleotide sequence of the full-length 30 protein coding sequence of clone AJ1\_1 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone AJ1\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone AJ1\_1 deposited under accession number ATCC 98278. In yet other preferred



embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 52 to amino acid 147.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1 or SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
- (b) the amino acid sequence of SEQ ID NO:2 from amino acid 52 to amino acid 147;
- (c) fragments of the amino acid sequence of SEQ ID NO:2; and
- (d) the amino acid sequence encoded by the cDNA insert of clone

AJ1\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 52 to amino acid 147.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4 from nucleotide 7 to nucleotide 1647;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4 from nucleotide 1 to nucleotide 305;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:5;

5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:5 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

10 (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:4 from nucleotide 7 to nucleotide 1647; the nucleotide sequence of SEQ ID NO:4 from  
15 nucleotide 1 to nucleotide 305; the nucleotide sequence of the full-length protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of  
20 clone AQ73\_3 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:5 from amino acid 1 to amino acid 68.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:4.

25 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:5;

30 (b) the amino acid sequence of SEQ ID NO:5 from amino acid 1 to amino acid 68;

(c) fragments of the amino acid sequence of SEQ ID NO:5; and

(d) the amino acid sequence encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:5 or the amino acid sequence of SEQ ID NO:5 from amino acid 1 to amino acid 68.

5 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 62 to nucleotide 757;
- 10 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 357 to nucleotide 703;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG142\_1 deposited under accession number ATCC 98278;
- 15 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG142\_1 deposited under accession number ATCC 98278;
- 20 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;
- 25 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- 30 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:6 from nucleotide 62 to nucleotide 757; the nucleotide sequence of SEQ ID NO:6 from nucleotide 357 to nucleotide 703; the nucleotide sequence of the full-length protein coding

sequence of clone BG142\_1 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone BG142\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7 from amino acid 184 to amino acid 214.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:6.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:7;
- (b) the amino acid sequence of SEQ ID NO:7 from amino acid 184 to amino acid 214;
- (c) fragments of the amino acid sequence of SEQ ID NO:7; and
- (d) the amino acid sequence encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:7 or the amino acid sequence of SEQ ID NO:7 from amino acid 184 to amino acid 214.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 404 to nucleotide 535;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1 to nucleotide 666;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV66\_1 deposited under accession number ATCC 98278;

- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV66\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV66\_1 deposited under accession number  
5 ATCC 98278;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV66\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;
- 10 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein  
15 of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:8 from nucleotide 404 to nucleotide 535; the nucleotide sequence of SEQ ID NO:8  
20 from nucleotide 1 to nucleotide 666; the nucleotide sequence of the full-length protein coding sequence of clone BV66\_1 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone BV66\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of  
25 clone BV66\_1 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 38.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:8.

30 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:9;

(b) the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 38;

(c) fragments of the amino acid sequence of SEQ ID NO:9; and

(d) the amino acid sequence encoded by the cDNA insert of clone  
5 BV66\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:9 or the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 38.

In one embodiment, the present invention provides a composition comprising an  
10 isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 1204 to nucleotide 1389;

15 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 881 to nucleotide 1380;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278;

20 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278;

25 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the  
30 amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:10 from nucleotide 1204 to nucleotide 1389; the nucleotide sequence of SEQ ID NO:10 from nucleotide 881 to nucleotide 1380; the nucleotide sequence of the full-length protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 59.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:10.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:11;
  - (b) the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 59;
  - (c) fragments of the amino acid sequence of SEQ ID NO:11; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:11 or the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 59.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 189 to nucleotide 1115;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 451;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:12 from nucleotide 189 to nucleotide 1115; the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 451; the nucleotide sequence of the full-length protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 88.



Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:12.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:13;
- (b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 88;
- (c) fragments of the amino acid sequence of SEQ ID NO:13; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:13 or the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 88.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 117 to nucleotide 923;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 174 to nucleotide 923;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 316;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CQ331\_2 deposited under accession number ATCC 98278;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CQ331\_2 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;

- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;
- 5 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions
- 10 to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 117 to nucleotide 923; the nucleotide sequence of SEQ ID NO:14 from nucleotide 174 to nucleotide 923; the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 316; the nucleotide sequence of the full-length protein coding

15 sequence of clone CQ331\_2 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone CQ331\_2 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278. In yet other preferred

20 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 57.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:14.

25 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:15;
- (b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to
- 30 amino acid 57;
- (c) fragments of the amino acid sequence of SEQ ID NO:15; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15 or the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 57.

In one embodiment, the present invention provides a composition comprising an  
5 isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 223 to nucleotide 483;
- 10 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 22 to nucleotide 397;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT550\_1 deposited under accession number ATCC 98278;
- 15 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT550\_1 deposited under accession number ATCC 98278;
- 20 (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;
- (i) a polynucleotide encoding a protein comprising a fragment of the  
25 amino acid sequence of SEQ ID NO:17 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- 30 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:16 from nucleotide 223 to nucleotide 483; the nucleotide sequence of SEQ ID NO:16 from nucleotide 22 to nucleotide 397; the nucleotide sequence of the full-length protein

coding sequence of clone CT550\_1 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone CT550\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 58.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:16.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:17;
  - (b) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 58;
  - (c) fragments of the amino acid sequence of SEQ ID NO:17; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:17 or the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 58.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 112 to nucleotide 969;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 154 to nucleotide 969;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 423;

- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278;
- 5 (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278;
- 10 (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity;
- 15 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- 20 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:18 from nucleotide 112 to nucleotide 969; the nucleotide sequence of SEQ ID NO:18 from nucleotide 154 to nucleotide 969; the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 423; the nucleotide sequence of the full-length protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278; or the

25 nucleotide sequence of the mature protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278. In yet other preferred

30 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 104.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:18.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:19;
- 5 (b) the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 104;
- (c) fragments of the amino acid sequence of SEQ ID NO:19; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;
- 10 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:19 or the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 104.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 15 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 37 to nucleotide 2766;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 243 to nucleotide 789;
- 20 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT797\_3 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278;
- 25 (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT797\_3 deposited under accession number ATCC 98278;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278;
- 30 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

5 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 37 to nucleotide 2766; the nucleotide sequence of SEQ ID NO:20 from nucleotide 243 to nucleotide 789; the nucleotide sequence of the full-length protein  
10 coding sequence of clone CT797\_3 deposited under accession number ATCC 98278; or the nucleotide sequence of the mature protein coding sequence of clone CT797\_3 deposited under accession number ATCC 98278. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278. In yet other preferred  
15 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 75 to amino acid 251.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

20 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 75 to  
25 amino acid 251;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins. Preferably such  
30 protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 75 to amino acid 251.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions.

Processes are also provided for producing a protein, which comprise:

- (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and
- (b) purifying the protein from the culture.

5 The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which  
10 specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

15

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

20

#### DETAILED DESCRIPTION

##### ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone  
25 in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have  
30 determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation



proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

5        Clone "AJ1\_1"

A polynucleotide of the present invention has been identified as clone "AJ1\_1". AJ1\_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer  
10 analysis of the amino acid sequence of the encoded protein. AJ1\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "AJ1\_1 protein").

The nucleotide sequence of the 5' portion of AJ1\_1 as presently determined is reported in SEQ ID NO:1. What applicants presently believe is the proper reading frame  
15 for the coding region is indicated in SEQ ID NO:2. The predicted amino acid sequence of the AJ1\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Additional nucleotide sequence from the 3' portion of AJ1\_1, including the polyA tail, is reported in SEQ ID NO:3.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone  
20 AJ1\_1 should be approximately 925 bp.

The predicted amino acid sequence disclosed herein for AJ1\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AJ1\_1 protein demonstrated at least some similarity to sequences identified as U39060 (GRIP1 [Mus musculus]). Based upon sequence similarity,  
25 AJ1\_1 proteins and each similar protein or peptide may share at least some activity.

Clone "AQ73\_3"

A polynucleotide of the present invention has been identified as clone "AQ73\_3". AQ73\_3 was isolated from a human adult ovary (PA-1 teratocarcinoma, untreated tissue  
30 pooled with retinoic-acid-treated and activin-treated tissue) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. AQ73\_3 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "AQ73\_3 protein").

The nucleotide sequence of AQ73\_3 as presently determined is reported in SEQ ID NO:4. What applicants presently believe to be the proper reading frame and the  
5 predicted amino acid sequence of the AQ73\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:5.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone AQ73\_3 should be approximately 2800 bp.

The nucleotide sequence disclosed herein for AQ73\_3 was searched against the  
10 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. AQ73\_3 demonstrated at least some similarity with sequences identified as AA514474 (nf57g01.s1 NCI\_CGAP\_Co3 Homo sapiens cDNA clone 924048), T47520 (Human hepatoma-derived growth factor (HDGF-2) cDNA), W24708 (zb62e08.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 308198 5'), and W45513  
15 (zc27g08.s1 Soares senescent fibroblasts NbHSF Homo sapiens cDNA clone 323582 3'). The predicted amino acid sequence disclosed herein for AQ73\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted AQ73\_3 protein demonstrated at least some similarity to sequences identified as D16431 (hepatoma-derived GF [Homo sapiens]), D63707 (mouse hepatoma  
20 derived growth factor (HDGF) [Mus musculus]), R66727 (Human hepatoma derived growth factor), U18997 (ORF\_f299 [Escherichia coli]), U97193 (similar to S. cerevisiae SIR2 (SP P06700) and mouse hepatoma derived growth factor HDGF (NID g945418) [Caenorhabditis elegans]), and W09404 (Human hepatoma-derived growth factor (HDGF-2)). Based upon sequence similarity, AQ73\_3 proteins and each similar protein  
25 or peptide may share at least some activity.

#### Clone "BG142\_1"

A polynucleotide of the present invention has been identified as clone "BG142\_1". BG142\_1 was isolated from a human adult brain cDNA library using methods which are  
30 selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BG142\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BG142\_1 protein").

The nucleotide sequence of BG142\_1 as presently determined is reported in SEQ ID NO:6. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BG142\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:7.

5 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BG142\_1 should be approximately 1100 bp.

The nucleotide sequence disclosed herein for BG142\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BG142\_1 demonstrated at least some similarity with sequences  
10 identified as AA170261 (ms87h11.r1 Soares mouse 3NbMS Mus musculus cDNA clone 618597 5' similar to TR E245601 E245601 G-RICH BOX-BINDING PROTEIN), L04282 (Human CACCC box-binding protein mRNA, complete cds), N27696 (yx51h12.r1 Homo sapiens cDNA clone 265319 5'), W96110 (ze09a11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 358460 5'), and W96111 (ze09a11.s1 Soares fetal heart NbHH19W  
15 Homo sapiens cDNA clone 358460 3'). The predicted amino acid sequence disclosed herein for BG142\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BG142\_1 protein demonstrated at least some similarity to sequences identified as U80078 (transcription factor BFCOL1 [Mus musculus]) and X98096 (G-rich box-binding protein [Mus  
20 musculus]). Based upon sequence similarity, BG142\_1 proteins and each similar protein or peptide may share at least some activity.

#### Clone "BV66\_1"

A polynucleotide of the present invention has been identified as clone "BV66\_1".  
25 BV66\_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BV66\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as  
30 "BV66\_1 protein").

The nucleotide sequence of BV66\_1 as presently determined is reported in SEQ ID NO:8. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BV66\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:9.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BV66\_1 should be approximately 870 bp.

The nucleotide sequence disclosed herein for BV66\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database. The nucleotide sequence of BV66\_1 indicates that it may contain a TAAA1 simple repeat element.

Clone "BV291\_3"

A polynucleotide of the present invention has been identified as clone "BV291\_3". BV291\_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. BV291\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "BV291\_3 protein").

The nucleotide sequence of BV291\_3 as presently determined is reported in SEQ ID NO:10. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the BV291\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:11.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone BV291\_3 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for BV291\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. BV291\_3 demonstrated at least some similarity with sequences identified as H10954 (ym06e09.r1 Homo sapiens cDNA clone 47034 5'), H10955 (ym06e09.s1 Homo sapiens cDNA clone 47034 3'), N25300 (yw52c10.s1 Homo sapiens cDNA clone 255858 3'), T25940 (Human gene signature HUMGS08173), T68890 (yc30g11.s1 Homo sapiens cDNA clone 82244 3'), T78286 (yc99a08.r1 Homo sapiens cDNA clone 24033 5'), Z39987 (H. sapiens partial cDNA sequence; clone c-1oh05), and Z47073 (Caenorhabditis elegans cosmid ZC506). The predicted amino acid sequence disclosed herein for BV291\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted BV291\_3 protein demonstrated at least some similarity to sequences identified as X02155 (BTTGR\_1 thyroglobulin [Bos taurus]). Based upon sequence similarity, BV291\_3 proteins and each

similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the BV291\_3 protein sequence centered around amino acid 48 of SEQ ID NO:11.

5        Clone "CK201\_1"

A polynucleotide of the present invention has been identified as clone "CK201\_1". CK201\_1 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer  
10 analysis of the amino acid sequence of the encoded protein. CK201\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CK201\_1 protein").

The nucleotide sequence of CK201\_1 as presently determined is reported in SEQ ID NO:12. What applicants presently believe to be the proper reading frame and the  
15 predicted amino acid sequence of the CK201\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:13.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CK201\_1 should be approximately 1080 bp.

The nucleotide sequence disclosed herein for CK201\_1 was searched against the  
20 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CK201\_1 demonstrated at least some similarity with sequences identified as AA129133 (zo09h12.s1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 567239 3' similar to contains Alu repetitive element), D81444 (Human fetal brain cDNA 5'-end GEN-164G10), R36326 (yg69h09.r1 Homo sapiens cDNA clone  
25 38821 5'), T08553 (Oncogene R-ras mutant cDNA (exons 2-6)), T31595 (Probe (BLUR13) for Alu repeat sequence), X03273 (Human Alu-family cluster 5' of alpha(1)-acid glycoprotein gene), and X69907 (H.sapiens gene for mitochondrial ATP synthase c subunit). The predicted amino acid sequence disclosed herein for CK201\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX  
30 search protocol. The predicted CK201\_1 protein demonstrated at least some similarity to sequences identified as D21827 (major surface glycoprotein [Pneumocystis carinii]). Based upon sequence similarity, CK201\_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CK201\_1 indicates that it may contain an Alu repetitive element.

Clone "CQ331\_2"

A polynucleotide of the present invention has been identified as clone "CQ331\_2". CQ331\_2 was isolated from a human adult heart cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
5 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CQ331\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CQ331\_2 protein").

The nucleotide sequence of CQ331\_2 as presently determined is reported in SEQ  
10 ID NO:14. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CQ331\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15. Amino acids 7 to 19 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 20, or are a transmembrane domain.

15 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CQ331\_2 should be approximately 1600 bp.

The nucleotide sequence disclosed herein for CQ331\_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CQ331\_2 demonstrated at least some similarity with sequences  
20 identified as J03766 (Canine cardiac calsequestrin mRNA, complete cds), L29766 (Homo sapiens epoxide hydrolase (EPHX) gene, complete cds), N83601 (KK1173F Homo sapiens cDNA clone KK1173 5' similar to CALSEQUESTRIN (CARDIAC)), T99646 (ye73f12.s1 Homo sapiens cDNA clone 123407 3' similar to contains Alu repetitive element;contains PTR5 repetitive element), and W76326 (zd60d04.r1 Soares fetal heart NbHH19W Homo  
25 sapiens cDNA clone 345031 5' similar to contains Alu repetitive element). The predicted amino acid sequence disclosed herein for CQ331\_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CQ331\_2 protein demonstrated at least some similarity to sequences identified as J03766 (DOGCAL\_1 Canine cardiac calsequestrin mRNA, complete cds [Canis canis])  
30 and X55040 (calsequestrin [Oryctolagus cuniculus]). Based upon sequence similarity, CQ331\_2 proteins and each similar protein or peptide may share at least some activity.

Clone "CT550\_1"

A polynucleotide of the present invention has been identified as clone "CT550\_1". CT550\_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CT550\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT550\_1 protein").

The nucleotide sequence of CT550\_1 as presently determined is reported in SEQ ID NO:16. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CT550\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT550\_1 should be approximately 1070 bp.

The nucleotide sequence disclosed herein for CT550\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database. The TopPredII computer program predicts a potential transmembrane domain within the CT550\_1 protein sequence centered around amino acid 25 of SEQ ID NO:17.

20

Clone "CT585\_1"

A polynucleotide of the present invention has been identified as clone "CT585\_1". CT585\_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CT585\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT585\_1 protein").

The nucleotide sequence of CT585\_1 as presently determined is reported in SEQ ID NO:18. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CT585\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19. Amino acids 2 to 14 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 15, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT585\_1 should be approximately 2710 bp.

The nucleotide sequence disclosed herein for CT585\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CT585\_1 demonstrated at least some similarity with sequences identified as AA069442 (zf74b02.s1 Soares pineal gland N3HPG Homo sapiens cDNA clone 382635 3'), L38961 (Homo sapiens putative transmembrane protein (B5) mRNA, complete cds), N34932 (yy49b10.s1 Homo sapiens cDNA clone 276859 3'), N60101 (TgESTzy11f10.r1 Toxoplasma gondii cDNA clone tgyz11f10.r1 5'), and U13019 (Caenorhabditis elegans cosmid T12A2). The predicted amino acid sequence disclosed herein for CT585\_1 was searched against the GenPept, GeneSeq, and SwissProt amino acid sequence databases using the BLASTX search protocol. The predicted CT585\_1 protein demonstrated at least some similarity to sequences identified as L34260 (transmembrane protein [Mus musculus]), L38961 (transmembrane protein [Homo sapiens]), P46975 (Caenorhabditis elegans oligosaccharyl transferase stt3 [Caenorhabditis elegans]), and U13019 (Caenorhabditis elegans cosmid T12A2 [Caenorhabditis elegans]). Based upon sequence similarity, CT585\_1 proteins and each similar protein or peptide may share at least some activity.

#### 20      Clone "CT797\_3"

A polynucleotide of the present invention has been identified as clone "CT797\_3". CT797\_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CT797\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CT797\_3 protein").

The nucleotide sequence of CT797\_3 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CT797\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CT797\_3 should be approximately 3300 bp.



The nucleotide sequence disclosed herein for CT797\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CT797\_3 demonstrated at least some similarity with sequences identified as AA573847 (nk08d06.s1 NCI\_CGAP\_Co2 Homo sapiens cDNA clone  
5 IMAGE:1012907). The predicted amino acid sequence disclosed herein for CT797\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CT797\_3 protein demonstrated at least some similarity to sequences identified as U18309 (chromokinesin [*Gallus gallus*]) and Z82271 (T01G1.1 [*Caenorhabditis elegans*]). Based upon sequence similarity, CT797\_3 proteins  
10 and each similar protein or peptide may share at least some activity.

#### Deposit of Clones

Clones AJ1\_1, AQ73\_3, BG142\_1, BV66\_1, BV291\_3, CK201\_1, CQ331\_2, CT550\_1, CT585\_1 and CT797\_3 were deposited on December 13, 1996 with the American Type  
15 Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98278, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

20 Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from  
25 pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse  
30 orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

	<u>Clone</u>	<u>Probe Sequence</u>
10	AJ1_1	SEQ ID NO:22
	AQ73_3	SEQ ID NO:23
	BG142_1	SEQ ID NO:24
	BV66_1	SEQ ID NO:25
	BV291_3	SEQ ID NO:26
15	CK201_1	SEQ ID NO:27
	CQ331_2	SEQ ID NO:28
	CT550_1	SEQ ID NO:29
	CT585_1	SEQ ID NO:30, SEQ ID NO:32
	CT797_3	SEQ ID NO:31

20

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
- (b) It should be designed to have a  $T_m$  of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

30

The oligonucleotide should preferably be labeled with  $\gamma$ -<sup>32</sup>P ATP (specific activity 6000 Ci/mmol) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated

label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately  $4 \times 10^6$  dpm/pmole.

5       The bacterial culture containing the pool of full-length clones should preferably be thawed and 100  $\mu$ l of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100  $\mu$ g/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the  
10       dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100  $\mu$ g/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

15       Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

20       The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100  $\mu$ g/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to  $1 \times 10^6$  dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The  
25       filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

30       The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S.

McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which the cDNA sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

	Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>‡</sup>	Hybridization Temperature and Buffer <sup>†</sup>	Wash Temperature and Buffer <sup>†</sup>
5	A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
	B	DNA:DNA	<50	T <sub>B</sub> <sup>*</sup> ; 1xSSC	T <sub>B</sub> <sup>*</sup> ; 1xSSC
	C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
	D	DNA:RNA	<50	T <sub>D</sub> <sup>*</sup> ; 1xSSC	T <sub>D</sub> <sup>*</sup> ; 1xSSC
	E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
	F	RNA:RNA	<50	T <sub>F</sub> <sup>*</sup> ; 1xSSC	T <sub>F</sub> <sup>*</sup> ; 1xSSC
10	G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
	H	DNA:DNA	<50	T <sub>H</sub> <sup>*</sup> ; 4xSSC	T <sub>H</sub> <sup>*</sup> ; 4xSSC
	I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
	J	DNA:RNA	<50	T <sub>J</sub> <sup>*</sup> ; 4xSSC	T <sub>J</sub> <sup>*</sup> ; 4xSSC
	K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
	L	RNA:RNA	<50	T <sub>L</sub> <sup>*</sup> ; 2xSSC	T <sub>L</sub> <sup>*</sup> ; 2xSSC
15	M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
	N	DNA:DNA	<50	T <sub>N</sub> <sup>*</sup> ; 6xSSC	T <sub>N</sub> <sup>*</sup> ; 6xSSC
	O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
	P	DNA:RNA	<50	T <sub>P</sub> <sup>*</sup> ; 6xSSC	T <sub>P</sub> <sup>*</sup> ; 6xSSC
	Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
	R	RNA:RNA	<50	T <sub>R</sub> <sup>*</sup> ; 4xSSC	T <sub>R</sub> <sup>*</sup> ; 4xSSC

<sup>‡</sup>: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

<sup>†</sup>: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

<sup>\*</sup>T<sub>B</sub> - T<sub>R</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(°C) = 81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds.,

5 John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or  
10 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an  
15 expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably  
20 linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the  
25 protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

30 Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant



methyle or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance  
5 with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

10 The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith,  
15 including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally  
20 provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another  
25 amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be  
30 expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

### USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

#### Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

#### Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

#### Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

- 5           The activity of a protein of the invention may, among other means, be measured by the following methods:

          Assays for T-cell or thymocyte proliferation include without limitation those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-  
10 Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bertagnolli et al., *J. Immunol.* 145:1706-1712, 1990; Bertagnolli et al., *Cellular Immunology* 133:327-341, 1991; Bertagnolli, et al., *J. Immunol.* 149:3778-3783, 1992; Bowman et al., *J. Immunol.* 152: 1756-1761, 1994.

- 15           Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon  $\gamma$ , Schreiber, R.D. In *Current Protocols in*  
20 *Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

- Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons,  
25 Toronto. 1991; deVries et al., *J. Exp. Med.* 173:1205-1211, 1991; Moreau et al., *Nature* 336:690-692, 1988; Greenberger et al., *Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human  
30 Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

#### Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosus in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient  
5 by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic  
10 acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function  
15 (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides.  
20 For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used  
25 to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II  
30 molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I  $\alpha$  chain protein and  $\beta_2$  microglobulin protein or an MHC class II  $\alpha$  chain protein and an MHC class II  $\beta$  chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.



Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

10       The activity of a protein of the invention may, among other means, be measured by the following methods:

      Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, *Immunologic studies in Humans*); Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Herrmann et al., *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann et al., *J. Immunol.* 128:1968-1974, 1982; Handa et al., *J. Immunol.* 135:1564-1572, 1985; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Bowman et al., *J. Virology* 61:1992-1998; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnoli et al., *Cellular Immunology* 133:327-341, 1991; Brown et al., *J. Immunol.* 153:3079-3092, 1994.

25       Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, *J. Immunol.* 144:3028-3033, 1990; and *Assays for B cell function: In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

30       Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

#### Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of  
5 hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or  
10 *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

15 Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood  
20 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of*  
25 *Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359,  
30 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

#### Tissue Growth Activity

5 A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

10 A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of  
15 congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce  
20 differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

25 Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and  
30 other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce  
5 differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in  
10 the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve  
15 tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present  
20 invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of  
25 non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac)  
30 and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

5 A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

10 Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium ).

15 Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

#### Activin/Inhibin Activity

20 A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin  $\alpha$  family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin- $\beta$  group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, 30 United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., *Endocrinology* 91:562-572, 1972; Ling et al., *Nature* 321:779-782, 1986; Vale et al., *Nature* 321:776-779, 1986; Mason et al., *Nature* 318:659-663, 1985; Forage et al., *Proc. Natl. Acad. Sci. USA* 83:3091-3095, 1986.

5

#### Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells.

10 Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses  
15 against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population  
20 of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent  
25 chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene  
30 Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. *J. Clin. Invest.* 95:1370-1376, 1995; Lind et al. *APMIS* 103:140-146, 1995; Muller et al. *Eur. J. Immunol.* 25: 1744-1748; Gruber et al. *J. of Immunol.* 152:5860-5867, 1994; Johnston et al. *J. of Immunol.* 153: 1762-1768, 1994.

#### Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

10       The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 15 35:467-474, 1988.

#### Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

30       The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and



Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 5 1995.

#### Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in 10 the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat 15 inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting 20 from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

#### Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major 25 roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

30 The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

5 E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and  
10 their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention  
15 encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue  
20 in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and  
25 polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the  
30 cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides  
5 encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

10 Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or  
15 tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

20 Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height,  
25 weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein,  
30 carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen  
5 in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

#### **ADMINISTRATION AND DOSING**

A protein of the present invention (from whatever source derived, including  
10 without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the  
15 effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem  
20 cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included  
25 in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers  
30 or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein

and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins  
5 including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other  
10 molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist  
15 in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S.  
20 Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or  
25 amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

30 In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines

or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If  
5 administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical  
10 composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is  
15 administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention.  
20 When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid  
25 form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present  
30 invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium

Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

5       The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician  
10 will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100  
15 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein of the present invention per kg body weight.

      The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is  
20 contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

      Protein of the invention may also be used to immunize animals to obtain  
25 polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in  
30 R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where

abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

5 For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably  
10 be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the  
15 methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical  
20 applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium  
25 sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other  
30 ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.



Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions  
5 from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of  
10 carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to  
15 provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in  
20 question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to  
25 humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of  
30 a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect

the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such  
5 polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present  
10 invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaVallie, Edward R.  
Racie, Lisa A.  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
Agostino, Michael J.
- (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 32
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Genetics Institute, Inc.
  - (B) STREET: 87 CambridgePark Drive
  - (C) CITY: Cambridge
  - (D) STATE: MA
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Sprunger, Suzanne A.
  - (B) REGISTRATION NUMBER: 41,323
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (617) 498-8284
  - (B) TELEFAX: (617) 876-5851

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 462 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

GTGGAAGGAG TGGATAATAA AATGAGTCAG TGCACCAGCT CCACCATTCC TAGCTCAAGT      60
CAAGAGAAAAG ACCCTAAAAT TAAGACAGAG ACAAGTGAAG AGGGATCTGG AGACTTGGAT      120
AATCTAGATG CTATTCTTGG TGATCTGACT AGTTCTGACT TTTACAATAA TTCCATATCC      180
TCAAATGGTA GTCATCTGGG GACTAAGCAA CAGGTGTTTC AAGGAACTAA TTCTCTGGGT      240
TTGAAAAGTT CACAGTCTGT GCAGTCTATT CGTCCTCCAT ATAACCGAGC AGTGTCTCTG      300
GATAGCCCTG TTTCTGTTGG CTCAAGTCCT CCAGTAAAAA ATATCAGTGC TTTCCCCATG      360
TTACCAAAGC AACCCATGTT GGGTGGGAAT CCAAGAATGA TGGATAGTCA RGAATAATTAT      420
GGCTCAAGTA TGGGAGACTG GGGCTTACCA AACTCAAAGG CC                          462

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ser Gln Cys Thr Ser Ser Thr Ile Pro Ser Ser Ser Gln Glu Lys
1           5           10           15
Asp Pro Lys Ile Lys Thr Glu Thr Ser Glu Glu Gly Ser Gly Asp Leu
20          25          30
Asp Asn Leu Asp Ala Ile Leu Gly Asp Leu Thr Ser Ser Asp Phe Tyr
35          40          45
Asn Asn Ser Ile Ser Ser Asn Gly Ser His Leu Gly Thr Lys Gln Gln
50          55          60
Val Phe Gln Gly Thr Asn Ser Leu Gly Leu Lys Ser Ser Gln Ser Val
65          70          75          80
Gln Ser Ile Arg Pro Pro Tyr Asn Arg Ala Val Ser Leu Asp Ser Pro
85          90          95

```

Val Ser Val Gly Ser Ser Pro Pro Val Lys Asn Ile Ser Ala Phe Pro  
 100 105 110

Met Leu Pro Lys Gln Pro Met Leu Gly Gly Asn Pro Arg Met Met Asp  
 115 120 125

Ser Gln Glu Asn Tyr Gly Ser Ser Met Gly Asp Trp Gly Leu Pro Asn  
 130 135 140

Ser Lys Ala  
 145

## (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 119 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 60

AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 119

## (2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3316 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGTAAGATGG CGGCTGTGAG TCTGCGGCTC GGCGACTTGG TGTGGGGGAA ACTCGGCCGA 60

TATCCTCCTT GGCCAGGAAA GATTGTTAAT CCACCAAAGG ACTTGAAGAA ACCTCGCGGA 120

AAGAAATGCT TCTTTGTGAA ATTTTGTGGA ACAGAAGATC ATGCCTGGAT CAAAGTGGAA 180

CAGCTGAAGC CATATCATGC TCATAAAGAG GAAATGATAA AAATTAACAA GGGTAAACGA 240

TTCCAGCAAG CGGTAGATGC TGTCGAAGAG TTCCTCAGGA GAGCCAAAGG GAAAGACCAG 300

ACGTCATCCC ACAATTCCTC TGATGACAAG AATCGACGTA ATTCCAGTGA GGAGAGAAGT	360
AGGCCAAACT CAGGTGATGA GAAGCGCAAA CTAGCCTGT CTGAAGGGAA GGTGAAGAAG	420
AACATGGGAG AAGGAAAGAA GAGGGTGTCT TCAGGCTCTT CAGAGAGAGG CTCCAAATCC	480
CCTCTGAAAA GAGCCCAAGA GCAAAGTCCC CGGAAGCGGG GTCGGCCCCC AAAGGATGAG	540
AAGGATCTCA CCATCCCGGA GTCTAGTACC GTGAAGGGGA TGATGGCCCG ACCGATGGCC	600
GCGTTTAAAT GGCAGCCAAC CGCAAGCGAG CCTGTTAAAG ATGCAGATCC TCATTTCCAT	660
CATTTCTGTC TAAGCCAAAC AGAGAAGCCA GCTGTCTGTT ACCAGGCAAT CACGAAGAAG	720
TTGAAAATAT GTGAAGAGGA AACTGGCTCC ACCTCCATCC AGGCAGCTGA CAGCACAGCC	780
GTGAATGGCA GCATCACACC CACAGACAAA AAGATAGGAT TTTTGGGCCCT TGGTCTCATG	840
GGAAGTGGA TCGTCTCCAA CTTGCTAAAA ATGGGTCACA CAGTGACTGT CTGGAACCGC	900
ACTGCAGAGA AAGAGGGGGC CCGTCTGGGA AGAACCCCCG CTGAAGTCGT CTCAACCTGC	960
GACATCACTT TCGCCTGCGT GTCGGATCCC AAGGCGGCCA AGGACCTGGT GCTGGGCCCC	1020
AGTGGTGTGC TGCAAGGGAT CCGCCCTGGG AAGTGCTACG TGGACATGTC AACAGTGGAC	1080
GCTGACACCG TCACTGAGCT GGCCAGGTG ATTGTGTCCA GGGGGGGCG CTTTCTGGAA	1140
GCCCCGTCT CAGGGAATCA GCAGCTGTCT AATGACGGGA TGTGTTGAT CTTAGCGGCT	1200
GGAGACAGGG GCTTATATGA GGA CTGCAGC AGCTGCTTCC AGGCGATGGG GAAGACCTCC	1260
TTCTTCCTAG GTGAAGTGG CAATGCAGCC AAGATGATGC TGATCGTGAA CATGGTCCAA	1320
GGGAGCTTCA TGGCCACTAT TGCCGAGGGG CTGACCCTGG CCCAGGTGAC AGGCCAGTCC	1380
CAGCAGACAC TCTTGGACAT CCTCAATCAG GGACAGTTGG CCAGCATCTT CCTGGACCAG	1440
AAGTGCCAAA ATATCTTGCA AGGAACTTT AAGCCTGATT TCTACCTGAA ATACATTCAG	1500
AAGGATCTCC GCTTAGCCAT TGCCTGGGT GATGCGGTCA ACCATCCGAC TCCCATGGCA	1560
GCTGCAGCAA ATGAGGTGTA CAAAAGAGCC AAGGCGCTGG ACCAGTCTGA CAACGATATG	1620
TCCGCCGTGT ACCGAGCCTA CATACACTAA GCTGTCGACA CCCC GCCCTC ACCCTCCAA	1680
TCCCCCTCT GACCCCTCT TCCTCACATG GGTGCGGGG CCTGGGAGTT CATTCTGGAC	1740
CAGCCCACCT ATCTCCATTT CCTTTTATAC AGACTTTGAG ACTTGCCATC AGCACAGCAC	1800
ACAGCAGCAC CCTTCCCTG AGGCCGGTGG GGAGGGGACA AGTGTGAGCA GGATTTGGCGT	1860
GTGGGAAAGC TCTTGAGCTG GGC ACTGGCC CCCC GGACGA GGTGGCTGTG TGTTCACACA	1920
CACACACACA CACACACACA CACAGGCTCT CGCCCCAGGA TAGAAGCTGC CCAGAACTG	1980

CTGCCTGGCT TTTTCTTTC CGAGCTTGTC TTATCTCAAA CCCCTTCCAG TCAAGGAACT	2040
AGAATCAGCA ACGAGAGTTG GAAGCCTTCC CACAGCTTCC CCCAGAGCGA AGAGGCTGTA	2100
GTCATGTCCC CATCCCCAC TGGATTCCCT ACAAGGAGAG GCCTTGGGCC CAGATGAGCC	2160
AGTACAGACT CCAGACAGAG GGGCCCTTGG GGCCCTCCAA CCTCAGGTGA TGAGCTGAGA	2220
AAGATGTTCA CGTCTAAGCG TCCAGTGTGC ACCCAGCGCT CCATAGACGC CTTTGTGAAC	2280
TGAAAAGAGA CTGGCAGAGT CCCGAGAAGA TGGGGCCCTG GCTTTCAGG GAGTGCAGCA	2340
AGCAGCCGGC CTGCAGACCC AGCCTGACCA ACGATGAGCA TTCTTAGGC TCAGCTCTTG	2400
ATACGGAAAC GAGTGTCTTC ACTCCAGCCA GCATCATGGT CTTCCGTGCT TCCCGGGCCC	2460
GGGGTCTGTC GGGAGGGAAG AGAACTGGGC CTGACCTACC TGAAGTGAAGT GGCCCTCCGA	2520
GGTGGGTCTG GGACATCCTA GAGGCCCTAC ATTTGTCTT GGATAGGGGA CCGGGGGGGG	2580
CTTGGAATGT TGCAAAAAA AAAGTTACCC AAGGGATGTC AGTTTTTTAT CCCTCTGCAT	2640
GGGTTGGATT TTCCAAAATC ATAATTTGCA GAAGGAAGGC CAGCATTTAC GATGCAATAT	2700
GTAATTATAT ATAGGGTGGC CACACTAGGG CGGGGTCTT CCCCCCTCAC AGCTTTGGCC	2760
CCTTTCAGAG ATTAGAACT GGGTTAGAGG ATTGAGAGA ACGAGTGGGG GGAGGGCAGG	2820
GAAGATGCCT GTCGGGTTTT TAGCACAGTT CATTTCACTG GGATTTTGAA GCATTTCTGT	2880
CTGGACACAA AGCCTGTTCT AGTCCTGGCG GAACACACTG GGGGTGGGGG CGGGGAAGA	2940
TGCGGTAATG AAACCGGTTA GTCAATTTTG TCTTAATATT GTTGACAATT CTGTAAAGTT	3000
CCTTTTTATG AATATTTCTG TTTAAGCTAT TTCACCTTTC TTTTGAAATC CTTCCTTTT	3060
AAGGAGAAAA TGTGACACTT GTGAAAAAGC TTGTAAAGAA GCCCCTCCCT TTTTCTTTA	3120
AACCTTTAAA TGACAAATCT AGGTAATTAA GGTGTGAAT TTTTATTTTT GCTTTGTTTT	3180
TAATGAACAT TTGTCTTTCA GAATAGGATT GTGTGATAAT GTTTAAATGG CAAAAACAAA	3240
ACATGATTTT GTGCAATTAA CAAAGCTACT GCAAGAAAAA TAAAACACTT CTTGGTAACA	3300
CAAAAAAAA AAAAAA	3316

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 547 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Ala	Val	Ser	Leu	Arg	Leu	Gly	Asp	Leu	Val	Trp	Gly	Lys	Leu	1	5	10	15
Gly	Arg	Tyr	Pro	Pro	Trp	Pro	Gly	Lys	Ile	Val	Asn	Pro	Pro	Lys	Asp	20	25	30	
Leu	Lys	Lys	Pro	Arg	Gly	Lys	Lys	Cys	Phe	Phe	Val	Lys	Phe	Phe	Gly	35	40	45	
Thr	Glu	Asp	His	Ala	Trp	Ile	Lys	Val	Glu	Gln	Leu	Lys	Pro	Tyr	His	50	55	60	
Ala	His	Lys	Glu	Glu	Met	Ile	Lys	Ile	Asn	Lys	Gly	Lys	Arg	Phe	Gln	65	70	75	80
Gln	Ala	Val	Asp	Ala	Val	Glu	Glu	Phe	Leu	Arg	Arg	Ala	Lys	Gly	Lys	85	90	95	
Asp	Gln	Thr	Ser	Ser	His	Asn	Ser	Ser	Asp	Asp	Lys	Asn	Arg	Arg	Asn	100	105	110	
Ser	Ser	Glu	Glu	Arg	Ser	Arg	Pro	Asn	Ser	Gly	Asp	Glu	Lys	Arg	Lys	115	120	125	
Leu	Ser	Leu	Ser	Glu	Gly	Lys	Val	Lys	Lys	Asn	Met	Gly	Glu	Gly	Lys	130	135	140	
Lys	Arg	Val	Ser	Ser	Gly	Ser	Ser	Glu	Arg	Gly	Ser	Lys	Ser	Pro	Leu	145	150	155	160
Lys	Arg	Ala	Gln	Glu	Gln	Ser	Pro	Arg	Lys	Arg	Gly	Arg	Pro	Pro	Lys	165	170	175	
Asp	Glu	Lys	Asp	Leu	Thr	Ile	Pro	Glu	Ser	Ser	Thr	Val	Lys	Gly	Met	180	185	190	
Met	Ala	Gly	Pro	Met	Ala	Ala	Phe	Lys	Trp	Gln	Pro	Thr	Ala	Ser	Glu	195	200	205	
Pro	Val	Lys	Asp	Ala	Asp	Pro	His	Phe	His	His	Phe	Leu	Leu	Ser	Gln	210	215	220	
Thr	Glu	Lys	Pro	Ala	Val	Cys	Tyr	Gln	Ala	Ile	Thr	Lys	Lys	Leu	Lys	225	230	235	240
Ile	Cys	Glu	Glu	Glu	Thr	Gly	Ser	Thr	Ser	Ile	Gln	Ala	Ala	Asp	Ser	245	250	255	
Thr	Ala	Val	Asn	Gly	Ser	Ile	Thr	Pro	Thr	Asp	Lys	Lys	Ile	Gly	Phe	260	265	270	



Leu Gly Leu Gly Leu Met Gly Ser Gly Ile Val Ser Asn Leu Leu Lys  
 275 280 285  
 Met Gly His Thr Val Thr Val Trp Asn Arg Thr Ala Glu Lys Glu Gly  
 290 295 300  
 Ala Arg Leu Gly Arg Thr Pro Ala Glu Val Val Ser Thr Cys Asp Ile  
 305 310 315 320  
 Thr Phe Ala Cys Val Ser Asp Pro Lys Ala Ala Lys Asp Leu Val Leu  
 325 330 335  
 Gly Pro Ser Gly Val Leu Gln Gly Ile Arg Pro Gly Lys Cys Tyr Val  
 340 345 350  
 Asp Met Ser Thr Val Asp Ala Asp Thr Val Thr Glu Leu Ala Gln Val  
 355 360 365  
 Ile Val Ser Arg Gly Gly Arg Phe Leu Glu Ala Pro Val Ser Gly Asn  
 370 375 380  
 Gln Gln Leu Ser Asn Asp Gly Met Leu Val Ile Leu Ala Ala Gly Asp  
 385 390 395 400  
 Arg Gly Leu Tyr Glu Asp Cys Ser Ser Cys Phe Gln Ala Met Gly Lys  
 405 410 415  
 Thr Ser Phe Phe Leu Gly Glu Val Gly Asn Ala Ala Lys Met Met Leu  
 420 425 430  
 Ile Val Asn Met Val Gln Gly Ser Phe Met Ala Thr Ile Ala Glu Gly  
 435 440 445  
 Leu Thr Leu Ala Gln Val Thr Gly Gln Ser Gln Gln Thr Leu Leu Asp  
 450 455 460  
 Ile Leu Asn Gln Gly Gln Leu Ala Ser Ile Phe Leu Asp Gln Lys Cys  
 465 470 475 480  
 Gln Asn Ile Leu Gln Gly Asn Phe Lys Pro Asp Phe Tyr Leu Lys Tyr  
 485 490 495  
 Ile Gln Lys Asp Leu Arg Leu Ala Ile Ala Leu Gly Asp Ala Val Asn  
 500 505 510  
 His Pro Thr Pro Met Ala Ala Ala Ala Asn Glu Val Tyr Lys Arg Ala  
 515 520 525  
 Lys Ala Leu Asp Gln Ser Asp Asn Asp Met Ser Ala Val Tyr Arg Ala  
 530 535 540  
 Tyr Ile His  
 545

(2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1097 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

ATTGGTTCAA GAAGAAAATT TGAGCCCAGG CACCCAAACA CCTTCAAATG ATAAAGCAAG      60
TATGTTGCAA GAATACTCCA AATACCTCCA ACAGGCTTTT GAAAAATCCA CTAATGCAAG      120
TTTACTCTT  GGACACGGTT TCCAATTTGT CAGTTTGTCT TCACCTCTCC ACAACCACAC      180
TTTGTTTCCA GAAAAACAAA TATACACTAC GTCTCCTTTG GAGTGTGGTT TCGGCCAATC      240
TGTTACCTCA GTGTTGCCAT CTTCAATTGCC AAAGCCTCCT TTTGGGATGT TGTTTGGATC      300
TCAGCCAGGT CTTTATTTGT CTGCTTTGGA TGCTACACAT CAGCAGTTGA CACCTTCCCA      360
GGAGCTGGAT GATCTGATAG ATTCTCAGAA GAACTTAGAG ACTTCATCAG CCTTCCAGTC      420
CTCATCTCAG AAATTGACTA GCCAGAAGGA ACAGAAAAAC TTAGAGTCTT CAACAGGCTT      480
TCAGATTCCA TCTCAGGAGT TAGCTAGCCA GATAGATCCT CAGAAAGACA TAGAGCCTAG      540
AACACGTAT  CAGATTGAGA ACTTTGCACA AGCGTTTGGT TCTCAGTTTA AGTCGGGCGA      600
CAGGGTGCCA ATGACCTTTA TCACTAACTC TAATGGAGAA GTGGACCATA GAGTAAGGAC      660
TTCAGTGTC  GATTTCTCAG GGTATACAAA TATGATGTCT GATGTAAGTG AGCCATGTAG      720
TACAAGAGTA AAGACACCCA CCAGCCAGAG TTACAGGTAA GGTCCCAAAA GTGGCCAGGC      780
TGGAGGTTTT TTAATGTAAT TTTGTTTAT  TTTGAGAACA CTGCCATTGG AATGTTTTTA      840
CACGATCCTA TTAAGAATAA TGTGATGCCC TTTCAATGCA ACTTTTCATA TTTAGTTTAT      900
TTTGTTAGCG TGATTTTAGC TCTGTTTGTA TTATGATTTT TAATCAAAAT CAATAGATTA      960
AAAATAGTTT GACATTCAAA GTGACAATGT TTAGCAATCA AATTTACATG TATAGATTGT      1020
CAGGGAATAG CCCAAATGTT TTAAACGCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA      1080
AAAAAAAAAA AAAAAAA                                     1097

```

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 232 amino acids

(B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Met Leu Gln Glu Tyr Ser Lys Tyr Leu Gln Gln Ala Phe Glu Lys Ser
 1             5             10             15

Thr Asn Ala Ser Phe Thr Leu Gly His Gly Phe Gln Phe Val Ser Leu
 20             25             30

Ser Ser Pro Leu His Asn His Thr Leu Phe Pro Glu Lys Gln Ile Tyr
 35             40             45

Thr Thr Ser Pro Leu Glu Cys Gly Phe Gly Gln Ser Val Thr Ser Val
 50             55             60

Leu Pro Ser Ser Leu Pro Lys Pro Pro Phe Gly Met Leu Phe Gly Ser
 65             70             75             80

Gln Pro Gly Leu Tyr Leu Ser Ala Leu Asp Ala Thr His Gln Gln Leu
 85             90             95

Thr Pro Ser Gln Glu Leu Asp Asp Leu Ile Asp Ser Gln Lys Asn Leu
100             105             110

Glu Thr Ser Ser Ala Phe Gln Ser Ser Ser Gln Lys Leu Thr Ser Gln
115             120             125

Lys Glu Gln Lys Asn Leu Glu Ser Ser Thr Gly Phe Gln Ile Pro Ser
130             135             140

Gln Glu Leu Ala Ser Gln Ile Asp Pro Gln Lys Asp Ile Glu Pro Arg
145             150             155             160

Thr Thr Tyr Gln Ile Glu Asn Phe Ala Gln Ala Phe Gly Ser Gln Phe
165             170             175

Lys Ser Gly Ser Arg Val Pro Met Thr Phe Ile Thr Asn Ser Asn Gly
180             185             190

Glu Val Asp His Arg Val Arg Thr Ser Val Ser Asp Phe Ser Gly Tyr
195             200             205

Thr Asn Met Met Ser Asp Val Ser Glu Pro Cys Ser Thr Arg Val Lys
210             215             220

Thr Pro Thr Ser Gln Ser Tyr Arg
225             230

```

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 775 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTGTCACATA CCACTCTTGT AGGTGTCCTC AATAATCCCC TTTTCCCACA AAATACACAG	60
GGTGTATTAT CTTTCTCTTT ATTCACCCCC ACTTTGCTGA ACTGAAGTTA ATTACATAGC	120
CTTTCTTCTA ACCTCCTTAG TAATGAACCT TCACATAAAG TGTATTTACA GCGTCTGTGG	180
TAGCCAGCCC TTCTCTCTCT ACTTTCTAGG AGGGGATAGC CAATAACTAG GAATTTAATG	240
ACAGATTTTT TTTTTCTTTG AAATAAATGG CCAGAGTTTC TCCATTTTAG AATTTTGTTG	300
TCCTCCTTAA TCATCTGCTT ACCTAGTCAT TACTCAATCT GCAGAACTT CATAAAGGAA	360
AAGTGCTGCA TTGTTTTTAC AAATAACAGT TTGTAGGGAA AATATGACAA ACCTCAACTA	420
TGGGAGTTGT CCACAATACA AAATTTTGAA AAAACATTAC ATAGTGATAA TATCATACTT	480
GGTTGTTAGG CTGTGTGCTT CCCACATCA GAGGCATCTA ATGATTATC TTTTGTAATT	540
GCTGTGAAC TTTTAAATA AGCCATTTAG TGTGAAATTG TCATGTATCA AATGGCTATT	600
GGAAATGGAC TTTACTCAAT TTTAATTCCA CTGCACTCTA GCCGGAGTGA CAGAGTAAGA	660
CTCTGTCTCA AAAATAAATA AATAAATAA TAAATAAATA AATAAATAA TAAATAAATA	720
ATAATAATAC AAGTTTTTCAT AAGGAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA	775

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 44 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Thr Asn Leu Asn Tyr Gly Ser Cys Pro Gln Tyr Lys Ile Leu Lys  
 1 5 10 15  
 Lys His Tyr Ile Val Ile Ile Ser Tyr Leu Val Val Arg Leu Val Ala  
 20 25 30  
 Ser Pro His Gln Arg His Leu Met Ile Tyr Leu Leu  
 35 40

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAAAAAGAAG ACAAAGCTCA CCTTCAGGCG GAGGTGCAGC ACCTGCGAGA GGACAACCTG 60  
 AGGCTACAGG AGGAGTCCCA GAACGCCTCG GACAAGCTGA AGAAGTTCAC AGAATGGGTC 120  
 TTCAACACCA TAGACATGAG CTAGGGAAGG CTGAGGAGGA CAGGAGAAGG GCCCAGACAC 180  
 TCCCTCCAGT GAGTGTCTTG CAGCCCTTAT TCCCTCCATA GAAAGCATCC TCAGAGCACC 240  
 TTCCCTGGCT TCCTACTCTG CCCCCTTTTG GGGAGTGCAC AACACAATAG TTGCAGATCA 300  
 ACAATCATCA CCTGCCTTTT GTAGAAAAGA AAAACAAAAA AAGTAAATAA AAATTTTAAA 360  
 CAGTAAATAA AAAGTTTAAC TGCTAAAATG TGAATGTCTT TATTTTTTTG CACAATATCT 420  
 TTATCTGTTA TGTATTTAAG AAGAACTGG GCCTTGGACC AGGGCGCCCC CTGGCCCATC 480  
 CGCCTCTATT CCCATCAGCT TTCTTATCAA CTTCAGGTAA CCCAAGCTTT CCCTTGTAT 540  
 TCTAACAAAT ATCATTATTC CTAGAAAAAG AATGTTTTTA TAACTTGTTT GGGGAGTAGA 600  
 GAGGGATATT TCCTTACCTT CTTCCCTAAA ATGCCTGGAG AGGGAGTTGC TTTGAGAAAA 660  
 TGCCTACCTC CCTTGAATGA CTCGTGCATG AGCTAGTGCT GTCTGTACCT GTCCTCCAGA 720  
 GATCAGCAGG ACCGGAGTTA AATATTTAAC AGCAAGTCTG TAAACCAGAG CAGCTCTGAC 780  
 AGTGCCTGCA GGCCACACCC CTTCTCAGTC CTGCATTGTG AGGTCAATTC CTGCTTCTCC 840  
 CTTTCCCCAG GAAGATGGTC CCACTTGTGC TGCAAGACTC TTTTTTGTIT GGCTTAATTG 900  
 AGCCCCACTA AATTGGAATC AATCTCTCTT ACAGCTTCCT GGCTCCAAAC ATTAATTGAT 960

```

TTCAGAATTC CCCCAAAC TA AAACCTTATC TGTCTGCATT TTGAATGCAT TTTGGTCAAA 1020
AGTATACGTT TTAAAGATTT TTAAAGATAA AAATGTGGCA CAACTGGTTT TTTTAGCTTG 1080
CTGAAAATGA CCATATCTCT AAATTAATCT TTCTCTCCAG AGCAAGACTT CACCAGTATT 1140
TGTAAC TAGG AGAAGCTAAG TGAATGTTTA ATTGTGAATT TTAATCATTG CTGTGTTAGGA 1200
ATAATGACTG TGATACTAGA ATGGGCTTTT GAAACCTGCA TGTCCTCAGT TGAAATTTCA 1260
GCACGGCATT TTCTGCATCC TTTCATGGCC ATCCAAAGGA TTCCGCTGCA GAAATTATTG 1320
ATGTGCTATT TTTGCTGTCT TGTGATGCAG GCTGCTTTGG GCCCCTGGGT CACTCTTCCA 1380
AGGCTGCTGT AGAGCACAGA GACATGGGGC TGGCCAGTGT TGA CTGACCT GAGGAGACCC 1440
CTTTGTTTTGT TGCCTTCATA ACTGTCACTA AACCGACCCC TCTGCCCTTT CAGTGGCAAC 1500
TCTGGTCTAA GGAACATTC AGCACTCTAG CGGCATCTGA TTGGAAGTTC CCTCACCCTAA 1560
GTAATCTCAA TTCCTTCCTC TCTCCATCCC TGAAAGAAAC AGGATGGATT TTCCTCTCTT 1620
CTCCCTGCTA CATTCACTAC CAGATTTTTA TGCTACAGTT TCATTCTTGA TTGTGATTTC 1680
TCCATGGAAT TTTTTTTTTT TGGTGACATT TCTATCATGG AAATAGGAAG ATTTCCGAGT 1740
GCTTTGTGAA GATTTCAATT GTCTGTCTCT TTCTCTCTTT GACTTGTATG AAGGAGATTG 1800
TACATTGCCT GATATCTCTT TGTAATGAG AAATATTGCT AACATCCAAG CATTCTGAAG 1860
TCTTGCTTAT CCTTCTGAGT TTAGTTCTCA TTTTGTTTTA CATTTTGT TT GGGGACTTGG 1920
GGCAAGCTAT TTATTAGAGT TTTGCAACAG AGTTCTTGTT TGAAGCCTCT AAAGACTACC 1980
TGTAATATTC AAAGAATAAA ATTCATTTTA AACGCTCTTT TAAAAAAAAA AAAAAAAAAA 2040
AAAAAAAAA AAAAAAAAAA 2060

```

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 62 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

Met Thr Val Ile Leu Glu Trp Ala Phe Glu Thr Cys Met Ser Gln Cys
1           5           10           15

```

Glu Ile Ser Ala Arg His Phe Leu His Pro Phe Met Ala Ile Gln Arg  
                   20                                  25                                  30

Ile Pro Leu Gln Lys Leu Leu Met Cys Tyr Phe Cys Cys Leu Val Met  
                   35                                  40                                  45

Gln Ala Ala Leu Gly Pro Trp Val Thr Leu Pro Arg Leu Leu  
                   50                                  55                                  60

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GATTATTTTC AGTAGGCAGA CATCTAATCG GAATCTTGCT CTGTGTGCCC AGGCTGGAGT 60

GTAATGGCAC AATCTCGGCT TACTGCAACC TCTGCCTCCT GGATTCAAGT GATTCTCCTG 120

CCTCAGCCTC CCAAGTAGCT GGGATTACAG CCCTGAAAAC CACTCGCTTG CAGAGCGCTG 180

GATCAGCAAT GCCTACTAGT TCTTCATTCA AACACCGGAT TAAAGAGCAG GAAGACTACA 240

TCCGAGATTG GACTGCTCAT CGAGAAGAGA TAGCCAGGAT CAGCCAAGAT CTTGCTCTCA 300

TTGCTCGGGA GATCAACGAT GTAGCAGGAG AGATAGATTG AGTGACTTCA TCAGGCACTG 360

CCCCTAGTAC CACAGTAAGC ACTGCTGCCA CCACCCCTGG CTCTGCCATA GACACTAGAG 420

AAGAGTTGGT TGATCGTGTT TTTGATGAAA GCCTCAACTT CCAAAAGATT CCTCCATTAG 480

TTCATTCCAA AACACCAGAA GGAAACAACG GTCGATCTGG TGATCCAAGA CCTCAAGCAG 540

CAGAGCCTCC CGATCACTTA ACAATTACAA GGCGGAGAAC CTGGAGCAGG GATGAAGTCA 600

TGGGAGATAA TCTGCTGCTG TCATCCGTCT TTCAGTTCTC TARGAAGATA AGACAATCTA 660

TAGATAAGAC AGCTGGAAAG ATCAGAATAT TATTTAAAGA CAAAGATCGG AATTGGGATG 720

ACATAGAAAG CAAATTAAGA GCCGAAAGTG AAGTCCCTAT TGTGAAAACC TCGAGCATGG 780

AGATTTCTTC TATCTTACAG GAACTGAAAA GAGTAGAAAA GCAGCTACAA GCAATCAATG 840

CTATGATTGA TCCTGATGGA ACTTTGGAGG CTCTGAACAA CATGGGATTT CCCAGTGCTA 900

TGTTGCCATC TCCACCGAAA CAGAAGTCCA GCCCTGTGAA TAACCACCAC AGCCCGGGTC 960

AGACACCAAC ACTTGCCAA CCAGAAGCTA GGGCTCTTCA TCCTGCTGCT GTTTCAGCCG 1020  
 CAGCTGAATT TGAGAATGCT GAATCTGAGG CTGATTTTCAG TATACATTTTC AATAGAGTCA 1080  
 ACCCTGATGG GGAAGAGGAA GATGTTACAG TAACATAAAT GACTTTCTCT TGATTGTTGA 1140  
 AAAAAAAAAA AAAAAAAAAA 1160

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 309 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Pro	Thr	Ser	Ser	Ser	Phe	Lys	His	Arg	Ile	Lys	Glu	Gln	Glu	Asp	1	5	10	15
Tyr	Ile	Arg	Asp	Trp	Thr	Ala	His	Arg	Glu	Glu	Ile	Ala	Arg	Ile	Ser	20	25	30	
Gln	Asp	Leu	Ala	Leu	Ile	Ala	Arg	Glu	Ile	Asn	Asp	Val	Ala	Gly	Glu	35	40	45	
Ile	Asp	Ser	Val	Thr	Ser	Ser	Gly	Thr	Ala	Pro	Ser	Thr	Thr	Val	Ser	50	55	60	
Thr	Ala	Ala	Thr	Thr	Pro	Gly	Ser	Ala	Ile	Asp	Thr	Arg	Glu	Glu	Leu	65	70	75	80
Val	Asp	Arg	Val	Phe	Asp	Glu	Ser	Leu	Asn	Phe	Gln	Lys	Ile	Pro	Pro	85	90	95	
Leu	Val	His	Ser	Lys	Thr	Pro	Glu	Gly	Asn	Asn	Gly	Arg	Ser	Gly	Asp	100	105	110	
Pro	Arg	Pro	Gln	Ala	Ala	Glu	Pro	Pro	Asp	His	Leu	Thr	Ile	Thr	Arg	115	120	125	
Arg	Arg	Thr	Trp	Ser	Arg	Asp	Glu	Val	Met	Gly	Asp	Asn	Leu	Leu	Leu	130	135	140	
Ser	Ser	Val	Phe	Gln	Phe	Ser	Xaa	Lys	Ile	Arg	Gln	Ser	Ile	Asp	Lys	145	150	155	160
Thr	Ala	Gly	Lys	Ile	Arg	Ile	Leu	Phe	Lys	Asp	Lys	Asp	Arg	Asn	Trp	165	170	175	



Asp Asp Ile Glu Ser Lys Leu Arg Ala Glu Ser Glu Val Pro Ile Val  
 180 185 190  
 Lys Thr Ser Ser Met Glu Ile Ser Ser Ile Leu Gln Glu Leu Lys Arg  
 195 200 205  
 Val Glu Lys Gln Leu Gln Ala Ile Asn Ala Met Ile Asp Pro Asp Gly  
 210 215 220  
 Thr Leu Glu Ala Leu Asn Asn Met Gly Phe Pro Ser Ala Met Leu Pro  
 225 230 235 240  
 Ser Pro Pro Lys Gln Lys Ser Ser Pro Val Asn Asn His His Ser Pro  
 245 250 255  
 Gly Gln Thr Pro Thr Leu Gly Gln Pro Glu Ala Arg Ala Leu His Pro  
 260 265 270  
 Ala Ala Val Ser Ala Ala Ala Glu Phe Glu Asn Ala Glu Ser Glu Ala  
 275 280 285  
 Asp Phe Ser Ile His Phe Asn Arg Val Asn Pro Asp Gly Glu Glu Glu  
 290 295 300  
 Asp Val Thr Val Thr  
 305

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1536 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GAAGAGAGAA AATCAGCCTG TCTGCTCTCT CTTGGCTCA ACAAGGCCTC TAACAGTCTT 60  
 CTGTCCTCTA TTCTGCACAC GGCATATTTG GGAACGAGAA ACAAAGTTT TCCCAAATGA 120  
 AGAGAACTCA CTGTGTTATT GTGGGGATTT ATTTTCTGTC CTCTGCAGG GCAGAAGAGG 180  
 GGCTTAATTT CCCCACATAT GATGGGAAGG ACCGAGTGGT AAGTCTTTCC GAGAAGAACT 240  
 TCAAGCAGGT TTAAAGAAA TATGACTTGC TTGCTCTA CTACCATGAG CCGGTGTCTT 300  
 CAGATAAGGT CACGCANAAA CAGTTCCAAC TGAAAGAAAT CGTGCTTGAG CTTGTGGCCC 360  
 ACGTCCTTGA ACATAAAGCT ATAGGCTTTG TGATGGTGA TGCCAAGAAA GAAGCCAAGC 420

```

TTGCCAAGAA ACTGGGTTTT GATGAAGAAG GAAGCCTGTA TATTCCTAAG GGTGATCGCA      480
CAATAGAGTT TGATGGCGAG TTTGCAGCTG ATGTCTTGGT GGAGTTCCTC TTGGATCTAA      540
TTGAAGACCC AGTGGAGATC ATCAGCAGCA AACTGGAAGT CCAAGCCTTC GAACGCATTG      600
AAGACTACAT CAAACTCATT GGCTTTTTCA AGAGTGAGGA CTCAGAATAC TACAAGGCTT      660
TTGAAGAAGC AGCTGAACAC TTCCAGCCTT ACATCAAATT CTTTGCCACC TTTGACAAAG      720
GGGTTGCAAA GAAATTATCT TTGAAGATGA ATGAGGTTGA CTTCTATGAG CCATTTATGG      780
ATGAGCCCAT TGCCATCCCC AACAAACCTT ACACAGAAGA GGAGCTGGTG GAGTTTGTGA      840
AGGAACACCA AAGGTGCCTG AGATGGCATG TGGGGGCTGG GGGCCTGGGG TCTGGGGAAT      900
GGAGAGGAGC CTCTCTGTGC TAACATTTCA GACCTGCCAA GAGCAACAAC CTAGTTAGTA      960
CCCCAGCAGT ACAGAACTCA GTAGTATGGC TTTGTTGATC AGTAATGACT AGCAGGGATG     1020
TTATTACTTC TGAATCTAAG TCTGCACCTG CAAGCAGAGT TTGATAAATC CCTCAGTCAG     1080
CAAATCCCCT CAAAGCCAGG GCAAGATATA AATAAAATTC TATACTAGGA ATGAGAGCAA     1140
TTTAGTGAAA GTTCCCATAT ACCAATAACC ATGCCCAGTG CTTTAGGGAA ACTATTTTAT     1200
CTAATCTCCA ACCTTAGGGA GTAATTATTA TTATCCCAAT TTTACAGATC AAGGAATTGG     1260
ACTCAATAGT TAAGTAACTT AGCCAAGGAT GAACACTCTA TGCATAGAAC TTCTGGGAGA     1320
GAAATGCTTG ATACCACTTA GTGTAGCTCC AGCATGGATC AGCAAACCTT TTCTGTAAAG     1380
AACAAAATGG TAAATATTTT AGGTTCTGTG GGCCAGATGG CGTCTGTAGC AACTACTTAA     1440
CTGCGGCTGT GGCATGAAAG CAGCCATGGA TCATGTATAA ACAAATGGGT GTGGCTGTGT     1500
ACCAAGTAAA GTTTATTTAG GAAAAAAAAA AAAAAA      1536

```

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 268 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

Met Lys Arg Thr His Leu Phe Ile Val Gly Ile Tyr Phe Leu Ser Ser
1           5           10           15

```

Cys Arg Ala Glu Glu Gly Leu Asn Phe Pro Thr Tyr Asp Gly Lys Asp  
 20 25 30  
 Arg Val Val Ser Leu Ser Glu Lys Asn Phe Lys Gln Val Leu Lys Lys  
 35 40 45  
 Tyr Asp Leu Leu Cys Leu Tyr Tyr His Glu Pro Val Ser Ser Asp Lys  
 50 55 60  
 Val Thr Xaa Lys Gln Phe Gln Leu Lys Glu Ile Val Leu Glu Leu Val  
 65 70 75 80  
 Ala His Val Leu Glu His Lys Ala Ile Gly Phe Val Met Val Asp Ala  
 85 90 95  
 Lys Lys Glu Ala Lys Leu Ala Lys Lys Leu Gly Phe Asp Glu Glu Gly  
 100 105 110  
 Ser Leu Tyr Ile Leu Lys Gly Asp Arg Thr Ile Glu Phe Asp Gly Glu  
 115 120 125  
 Phe Ala Ala Asp Val Leu Val Glu Phe Leu Leu Asp Leu Ile Glu Asp  
 130 135 140  
 Pro Val Glu Ile Ile Ser Ser Lys Leu Glu Val Gln Ala Phe Glu Arg  
 145 150 155 160  
 Ile Glu Asp Tyr Ile Lys Leu Ile Gly Phe Phe Lys Ser Glu Asp Ser  
 165 170 175  
 Glu Tyr Tyr Lys Ala Phe Glu Glu Ala Ala Glu His Phe Gln Pro Tyr  
 180 185 190  
 Ile Lys Phe Phe Ala Thr Phe Asp Lys Gly Val Ala Lys Lys Leu Ser  
 195 200 205  
 Leu Lys Met Asn Glu Val Asp Phe Tyr Glu Pro Phe Met Asp Glu Pro  
 210 215 220  
 Ile Ala Ile Pro Asn Lys Pro Tyr Thr Glu Glu Glu Leu Val Glu Phe  
 225 230 235 240  
 Val Lys Glu His Gln Arg Cys Leu Arg Trp His Val Gly Ala Gly Gly  
 245 250 255  
 Leu Gly Ser Gly Glu Trp Arg Gly Ala Ser Leu Cys  
 260 265

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1009 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAATTCGGCC TTCATGCGCC TGCAGGAAAG AATCTGACAT CATCACACTG TGT'TTCCTT	60
AACTTGACAG GAAGTCAACT TCAAGCAGAT TGACTTGAAA CGGGATCTCA TTTAGGAAGC	120
ATAAGTGTC AATCAAAAAC TGTGTATTTT TTTAAATTG GAAAATACTC AAGTTCCAGT	180
TGCTTATCAT TCTCCTTCAC TTTCTGAAAA CCTGGCAATC CCATGTGGAC TTCTGGTAGA	240
ATGAGCAATG CAAAGAAC TGCTGGACTT GGCATGTCCT TGTACTTCTG GGGGCTGATG	300
GACCTTACGA CCACCGTTCT CTCGGACACC CCAACACCAC AAGGTGAATT AGAAGCACTC	360
CTGTCAGACA AGCCACAGTC ACATCAGCGG ACCAAGARGA GCTGGGTTTG GAACCAGTTT	420
TTCTGTCTGG AAGAGTACAC TGGGACCGAC CCTTTGTATG TCGGCAAGGT AAGAAATGCC	480
AAGTAGAAAT GACCCGGGTA GTGGATATTG AAATTGAATA TGAATTGAGT ATCAAAGTTG	540
ACCTAGCCTT TATYTGAGAC CTGAGAAAAA CTAGAACAAG TGGTACGTTA CTTGACACCT	600
AGCTAAAATG TAACTTCTGC TTTGTCAGAG ACCAGTCTGA AAGGAAAGAT TTATTTCTTT	660
GTCCATGTCT CTGGTATGAA TGGGAAAAAG TGGGAATTGG GATT'TGGAGG AAAAGGCTCA	720
GACCCTGCAA GAGCTATTCA AGTCCTAAAA GAGGCAGCAG CAGCTGTCTG GGAATGACAG	780
AATGGGGGAG AGGGAAACTT GGAAATACAA GAAGAGTACA GAG'TTT'TTG CTTTGTGTTT	840
TTATGGGGTT TTTTTCAGAG CATTTCCAGA GGTATGTCCT GAGGTGCAAA TTTAATGAAA	900
AAAATAAAAA TAAAACATTT TTCATTTTCA AAGATCTTAG CATGTGCTTT AGGATAGTTG	960
GAGACAATAA ATATATTTAT AAATGTTAAA AAAAAAAAAA AAAAAAAAAA	1009

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 87 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Trp Thr Ser Gly Arg Met Ser Asn Ala Lys Asn Trp Leu Gly Leu  
 1 5 10 15  
 Gly Met Ser Leu Tyr Phe Trp Gly Leu Met Asp Leu Thr Thr Thr Val  
 20 25 30  
 Leu Ser Asp Thr Pro Thr Pro Gln Gly Glu Leu Glu Ala Leu Leu Ser  
 35 40 45  
 Asp Lys Pro Gln Ser His Gln Arg Thr Lys Xaa Ser Trp Val Trp Asn  
 50 55 60  
 Gln Phe Phe Val Leu Glu Glu Tyr Thr Gly Thr Asp Pro Leu Tyr Val  
 65 70 75 80  
 Gly Lys Val Arg Asn Ala Lys  
 85

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2546 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AAAAGAAACC AAGGAAATTT GTATGATAAG GCAGGTAAG TGAGGAAACA TGCAACTGAA 60  
 CAGGAAAAAA CTGAAGAGGG ATTAGGCCCT AATATAAAAA GCATTGTCAC CATGTTGATG 120  
 CTGATGCTAT TGATGATGTT TGCTGTCCAC TGTACCTGGG TCACAAGCAA TGCCTACTCT 180  
 AGTCCAAGTG TAGTCCTGGC CTCATACAAT CATGATGGCA CCAGGAATAT CTTAGATGAT 240  
 TTTAGAGAAG CTTACTTTTG GCTAAGGCAA AATACAGATG AACATGCACG AGTAATGTCT 300  
 TGGTGGGATT ATGGCTATCA GATAGCTGGA ATGGCTAATA GAACTACGTT GGTGGATAAT 360  
 AACACCTGGA ATAACAGCCA CATAGCACTG GTGGGAAAAG CTATGTCTTC TAATGAAACA 420  
 GCAGCCTATA AAATCATGAG GACTCTAGAT GTAGATTATG TTTTGGTTAT TTTTGGAGGG 480  
 GTTATTGGCT ATTCTGGTGA TGATATCAAC AAATTTCTCT GGATGGTTAG GATAGCTGAA 540  
 GGAGAACATC CCAAAGACAT TCGGGAAAGT GACTATTTTA CCCCACAGGG AGAATTCGGT 600  
 GTAGACAAAG CAGGATCCCC TACTTTGTTG AATTGCCTTA TGTATAAAAT GTCATACTAC 660

AGATTGGAG	AAATGCAGCT	GGATTTTCGT	ACACCCCAG	GTTTGGACCG	AACACGTAAT	720
GCTGAGATTG	GAAATAAGGA	CATTAAATTC	AAACATTTGG	AAGAAGCCTT	TACATCAGAA	780
CAC'TGGCTTG	TTAGGATATA	TAAAGTAAAA	GCACCTGATA	ACAGGGAGAC	ATTAGATCAC	840
AAACCTCGAG	TCACCAACAT	TTTCCCAAAA	CAGAAGTATT	TGTCAAAGAA	GACTACCAAA	900
AGGAAGCGTG	GCTACATTAA	AAATAAGCTG	GTTTTTAAGA	AAGGCAAGAA	AATATCTAAG	960
AAGACTGTTT	AAATGCAC'TG	TTCTGGTTCC	TAAC'TTGAAG	CAGTTGTCC'T	TGTGAGAACC	1020
GGTCTTTGCC	TTTAGCTCAT	GTCGTGTTTC	ACAGCAAAGA	GGGTACAGAA	CCATCACTGG	1080
TCCAGGTTAA	TGTACAAAAT	TTTCTGGCAA	TGCCTGATTA	AAAAAATAAA	ATTGGCTTGT	1140
TGAGAACAGC	TGTTTTCGAT	TTCTAATGTG	AAGCAAGACA	GAGCACTGCT	GTAAATGTCT	1200
AGCAGCAGAT	TTTTTTTTTA	TTGGTACATA	TTATCCTTCA	AATCTGAGAA	TTTGGACTAA	1260
CTGCACCAAA	GAACCCTCTA	ATTTGGTCCC	TGGCACATGC	ATACTTGTCA	ATGTTTTTAT	1320
TCTCTTACAA	GACCTGCATT	TTATTTGAAT	TACCCGAATA	GCAATATGTA	AAATACAAGT	1380
GACAAAATGT	GATGAGAGCT	TCTTGAACCG	GTAAACTAGT	ACAGGTCTGA	GAAAGACATA	1440
TTAGAAGAAT	CATTATACTT	CCCTGAATTA	TATTTATTTT	CATGTTTCTC	TAATGCAAAG	1500
AATGTTTCAT	CAATGTATA	TTTTCTGTTG	CTTACTGTTT	GCTCTGAGAA	GAAGCTGCTG	1560
TTTCAAAGAT	GGACCTCTGA	GTAGCTAATT	GATTCAAGTA	GTTTTTTTAT	GTTGACACAT	1620
TATTACTGCT	GTTAGCAGTC	GTTTTACCA	GGTACTTACA	GAGCAGATTT	CATACATCAT	1680
TCATTCAAGG	GCTAAATTTA	TATTTTTTGG	AAATCATGGC	AACTACACAG	GATGTTGCTT	1740
ACCAGGACGG	AGTTTTGGTA	TCTTAGTACT	GAAGTTAGCA	CTATGTTTAC	ATGCAAAAGA	1800
TTAAGGAAAA	AACCCTTAAA	GTGGACAGGT	ATCCAAAGTT	CATTTTCTGT	GACTCATCAA	1860
AGTGACAAAA	GACTTGTAAC	AACTTTGCCT	GGACTTTTTT	CATTTTACAA	CAGTTCATCC	1920
ATTACAAATG	ATTTTGTCT	CTGCTCCATA	TTTTTTAATC	CCTTAAGCAT	TTGATGAAAC	1980
ACTCTTTAGT	GCTATATGCA	TTTCTTACT	TTTGTTAAAA	ATGTGACAAT	TGTCAAAAAA	2040
TGCAC'TAAAA	TGTAAATGGA	GATTGAACAA	GTTCACTTTC	CAGCTTATAG	GCAACTTTAT	2100
ACAGACTTGA	ACA'TTTTCTC	CAGTTGTTTA	GTAAAAGTGA	AAGAGAAAGG	GTTTTTCCTG	2160
CCACAGGATA	TAAC'TTTTTT	TTATATAACA	AGCATAACAC	ACCACTGCTT	TTGGTGAAAA	2220
AGTGCAGAAT	AGTATGTACC	TTTTATGAAG	AAAAATGTAA	TTTACAATAT	TCAGTGAGAA	2280
TGTTACTGCT	GATTTTCTTT	TCCAAGGTGT	AGAATATCTT	TTGATTTATA	GAATTCATTT	2340

TTGACCCAGA TGATGGTTC TTTACAGAAC AATAAAATGG CTGAACATTT TCACAAATAG 2400  
 AGTGTAACGA AGTCTGGATT TCTGATACCT TGTCATTTGG GGGATTTTAT TTTACTTTGT 2460  
 TGCTTTAAAA TTCAATGCAG AGAAGTTGTT GACTGTAGGG GAAATAAAGT TAATTCAAAT 2520  
 TTTGAAAAAA AAAAAAAAAA AAAAAA 2546

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 286 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Leu Met Leu Met Leu Leu Met Met Phe Ala Val His Cys Thr Trp  
 1 5 10 15  
 Val Thr Ser Asn Ala Tyr Ser Ser Pro Ser Val Val Leu Ala Ser Tyr  
 20 25 30  
 Asn His Asp Gly Thr Arg Asn Ile Leu Asp Asp Phe Arg Glu Ala Tyr  
 35 40 45  
 Phe Trp Leu Arg Gln Asn Thr Asp Glu His Ala Arg Val Met Ser Trp  
 50 55 60  
 Trp Asp Tyr Gly Tyr Gln Ile Ala Gly Met Ala Asn Arg Thr Thr Leu  
 65 70 75 80  
 Val Asp Asn Asn Thr Trp Asn Asn Ser His Ile Ala Leu Val Gly Lys  
 85 90 95  
 Ala Met Ser Ser Asn Glu Thr Ala Ala Tyr Lys Ile Met Arg Thr Leu  
 100 105 110  
 Asp Val Asp Tyr Val Leu Val Ile Phe Gly Gly Val Ile Gly Tyr Ser  
 115 120 125  
 Gly Asp Asp Ile Asn Lys Phe Leu Trp Met Val Arg Ile Ala Glu Gly  
 130 135 140  
 Glu His Pro Lys Asp Ile Arg Glu Ser Asp Tyr Phe Thr Pro Gln Gly  
 145 150 155 160  
 Glu Phe Arg Val Asp Lys Ala Gly Ser Pro Thr Leu Leu Asn Cys Leu  
 165 170 175

Met Tyr Lys Met Ser Tyr Tyr Arg Phe Gly Glu Met Gln Leu Asp Phe  
 180 185 190

Arg Thr Pro Pro Gly Phe Asp Arg Thr Arg Asn Ala Glu Ile Gly Asn  
 195 200 205

Lys Asp Ile Lys Phe Lys His Leu Glu Glu Ala Phe Thr Ser Glu His  
 210 215 220

Trp Leu Val Arg Ile Tyr Lys Val Lys Ala Pro Asp Asn Arg Glu Thr  
 225 230 235 240

Leu Asp His Lys Pro Arg Val Thr Asn Ile Phe Pro Lys Gln Lys Tyr  
 245 250 255

Leu Ser Lys Lys Thr Thr Lys Arg Lys Arg Gly Tyr Ile Lys Asn Lys  
 260 265 270

Leu Val Phe Lys Lys Gly Lys Lys Ile Ser Lys Lys Thr Val  
 275 280 285

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4061 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AAGCAATTGA AGAAATTGCA GCAGGATGTG ATGGAAATGA AAAAAACAAA GGTTCGCCTA 60

ATGAAACAAA TGAAAGAAGA ACAAGAGAAA GCCAGACTGA CTGAGTCTAG AAGAAACAGA 120

GAGATTGCTC AGTTGAAAAA GGATCAACGT AAAAGAGATC ATCMACCTAG ACTTCTGGAA 180

GCCCCAAAAA GAAACCAAGA AGTGGTTCTA CGTCGCAAAA CTGAAGAGGT TACGGCTCTT 240

CGTCGGCAAG TAAGACCCAT GTCAGATAAA GTGGCTGGGA AAGTTACTCG GAAGCTGAGT 300

TCATCTGATG CACCTGCTCA GGACACAGGT TCCAGTGCAG CTGCTGTCTG AACAGATGCA 360

TCAAGGACAG GAGCCCAGCA GAAAATGAGA ATTCTGTGG CGAGAGTCCA GGCCTTACCA 420

ACGCCGGCAA CAAATGGAAA CAGGAAAAAA TATCAGAGGA AAGGATTGAC TGGCCGAGTG 480

TTTATTTCCTA AGACAGCTCG CATGAAGTGG CAGCTCCTTG AGCGCAGGGT CACAGACATC 540

ATCATGCAGA AGATGACCAT TTCCAACATG GAGGCAGATA TGAATAGACT CCTCAAGCAA 600



CGGGAGGAAC TCACAAAAAG ACGAGAGAAA CTTTCAAAAA GAAGGGAGAA GATAGTCAAG	660
GAGAATGGAG AGGGAGATAA AAATGTGGCT AATATCAATG AAGAGATGGA GTCAGTACT	720
GCTAATATCG ATTACATCAA TGACAGTATT TCTGATTGTC AGGCCAACAT AATGCAGATG	780
GAAGAAGCAA AGGAAGAAGG TGAGACATTG GATGTTACTG CAGTCATTAA TGCCTGCACC	840
CTTACAGAAG CCCGATACCT GCTAGATCAC TTCTGTCAA TGGGCATCAA TAAGGGTCTT	900
CAGGCTGCCC AGAAAGAGGC TCAAATTAAA GTACTGGAAG GTCGACTCAA ACAAACAGAA	960
ATAACCAGTG CTACCCAAAA CCAGCTCTTA TTCCATATGT TGAAAGAGAA GGCAGAATTA	1020
AATCCTGAGC TAGATGCTTT ACTAGGCCAT GCTTTACAAG ATCTAGATAG CGTACCATTA	1080
GAAAATGTAG AGGATAGTAC TGATGAGGAT GCTCCTTTAA ACAGCCCAGG ATCAGAAGGA	1140
AGCACGCTGT CTTTCAGATCT CATGAAGCTT TGTGGTGAAG TGAAACCTAA GAACAAGGCC	1200
CGAAGGAGAA CCACCACTCA GATGGAATTG CTGTATGCAG ATAGCAGTGA ACTAGCTTCA	1260
GACACTAGTA CAGGAGATGC CTCCTTGCCT GGCCCTCTCA CACCTGTTGC AGAAGGGCAA	1320
GAGATTGGAA TGAATACAGA GACAAGTGGT ACTTCTGCTA GGGAAAAAGA GCTCTCTCCC	1380
CCACCTGGCT TACCTTCTAA GATAGGCAGC ATTTCCAGGC AGTCATCTCT ATCAGAAAAA	1440
AAAATTCCAG AGCCTTCTCC TGTAACAAGG AGAAAGGCAT ATGAGAAAGC AGAAAAATCA	1500
AAGGCCAAGG AACAAAAGCA CTCAGATTCT GGAACCTCAG AGGCTAGTCT TTCACCTCCT	1560
TCTTCCCCAC CAAGCCGGCC CCGTAATGAA CTGAATGTTT TTAATCGTCT TACTGTTTCT	1620
CAGGGAAACA CATCAGTTCA GCAGGATAAG TCTGATGAAA GTGACTCCTC TCTCTCGGAG	1680
GTACACAGCA GATCCTCCAG AAGGGGCATA ATCAACCCAT TTCCTGCTTC AAAAGGAATC	1740
AGAGCTTTTC CACTTCAGTG TATTACATA GCTGAAGGGC ATACAAAAGC TGTGCTCTGT	1800
GTGGATTCTA CTGATGATCT CCTCTTCACT GGATCAAAAG ATCGTACTTG TAAAGTATGG	1860
AATCTGGTGA CTGGGCAGGA AATAATGTCA CTGGGGGGTC ATCCCAACAA TGTCGTGTCT	1920
GTAAAATACT GTAATTATAC CAGTTTGGTC TTCACTGTAT CAACATCTTA TATTAAGGTG	1980
TGGGATATCA GAGATTGAGC AAAGTGCATT CGAACACTAA CGTCTTCAGG TCAAGTTACT	2040
CTTGGAGATG CTTGTTCTGC AAGTACCAGT CGAACAGTAG CTATTCCCTC TGGAGAGAAC	2100
CAGATCAATC AAATTGCCCT AAACCCAACT GGCACCTTCC TCTATGCTGC TTCTGGAAAT	2160
GCTGTCAGGA TGTGGGATCT TAAAAGGTTT CAGTCTACAG GAAAGTTAAC AGGACACCTA	2220
GGCCCTGTTA TGTGCCTTAC TGTGGATCAG ATTTCCAGTG GACAAGATCT AATCATCACT	2280

GGCTCCAAGG ATCATTACAT CAAAATGTTT GATGTTACAG AAGGAGCTCT TGGGACTGTG	2340
AGTCCCACCC ACAATTTTGA ACCCCCTCAT TATGATGGCA TAGAAGCACT AACCATTCAA	2400
GGGGATAACC TATTTAGTGG GTCTAGAGAT AATGGAATCA AGAAATGGGA CTTAACTCAA	2460
AAAGACCTTC TTCAGCAAGT TCCAAATGCA CATAAGGATT GGGTCTGTGC CCTGGGAGTG	2520
GTGCCAGACC ACCCAGTTTT GCTCAGTGGC TGCAGAGGGG GCATTTTGAA AGTCTGGAAC	2580
ATGGATACTT TTATGCCAGT GGGAGAGATG AAGGGTCATG ATAGTCCTAT CAATGCCATA	2640
TGTGTTAATT CCACCCACAT TTTTACTGCA GCTGATGATC GAACTGTGAG AATTGGAAG	2700
GCTCGCAATT TGCAAGATGG TCAGATCTCT GACACAGGAG ATCTGGGGGA AGATATTGCC	2760
AGTAATTAAA CATGAATGAA GATAGGTTGT AAAGTGAATG CTGTGATAAT ACTCTGTATT	2820
CTTTATGGAA AATGTTGTCC TGTACTTACT AGGCAAAACG TATGAATCGG ATTAAGTGA	2880
AAATATATCT GAATTCAACT GCTGACTATA AATGGTATTC TAATAAAATT GTGTACTATC	2940
CTGTGTGCTT AGTTTAAAGA TCAACCAATA GATATATATC CTACAATTGA TATATTGCTT	3000
TATTCACACT TTTATTGTGG CTGAATTTTT GTGCCTATCT ATAAACACA CTTTCAAATT	3060
ATTTGAATTA CCAAGACGTC TGCTTTGTG ACAGTCAGAA AACACACCTG GAATACGATG	3120
CAGCCCACCA TTAAGTCATT CATGTAGTTT ATTCAAGTGA TTTATGTATT TAAAGTAAAT	3180
ATTGAAATG TTAGTCAAAT TGTGGTTTGC TTGTCAGGTA TTTATATCAG TCTGTAGTGG	3240
ATTCCCAAAT TTCAAAGCTC TTTTAATGTA ATGGACAAAA ATAAGATATG AGAATATTAT	3300
TGATGAATTT TCATAAGGTG GAATTGATCT TAATCTACTA ACAGAGAAGG GTAGACAGTT	3360
TGTGTTAAAT GTTGGCATT ACCTGTATTG ACCAAAGTTT TGCAGCTCTA CTATATTCTG	3420
TGCTCAGGAC TAAAATGCTG TTAATTTTTT TTTTTTTTTT TCCAGTGCTG TGCATATATT	3480
CTGTGATGGG AAACATTGTT GATGTCCTAA CAGAAATATA TTTTGATCTA TTTTCCTATG	3540
GAGTTGTTTC TATTATGACC ATTTAATTTTT GTTTTATTTT AATAGTAGTA TTTCTTCCC	3600
TTTTATCTAA TTTTATATAT GCTGCTAAAT ATATTTTAAA TATACTATGT TTGCGAACCT	3660
TGGTAGCTAT GATGAGAGCT ATTATCATCT GTGGTGGGAA AAGCTATGTA AATAGGTAGA	3720
TTGTATAGAG AGACTATCTT GTGTGTGCC TGTATGAATT TTTAAAAGTT GTTGACTGGA	3780
TTTTGCAAAA GGATGTATAA TATTTCTGTC TGCTCAGAAT ATTAATTTGT AAATCTGCA	3840
AGTTTAATTT TTATGTAGAT GGTATAACAT TTGAAAATAT TGTCTTATGT GATTTTCTCC	3900
CCTGAAAATA TTTGCTTGTA AATGAAAAC TAGCTAGGGC TTAAATAAAC ATGTTGCTAT	3960

GAAATKRAAA AAAAAAAAAA AAAARAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 4020  
 AARAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA A 4061

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 910 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Lys	Lys	Thr	Lys	Val	Arg	Leu	Met	Lys	Gln	Met	Lys	Glu	Glu	Gln	1	5	10	15
Glu	Lys	Ala	Arg	Leu	Thr	Glu	Ser	Arg	Arg	Asn	Arg	Glu	Ile	Ala	Gln	20	25	30	
Leu	Lys	Lys	Asp	Gln	Arg	Lys	Arg	Asp	His	Xaa	Leu	Arg	Leu	Leu	Glu	35	40	45	
Ala	Gln	Lys	Arg	Asn	Gln	Glu	Val	Val	Leu	Arg	Arg	Lys	Thr	Glu	Glu	50	55	60	
Val	Thr	Ala	Leu	Arg	Arg	Gln	Val	Arg	Pro	Met	Ser	Asp	Lys	Val	Ala	65	70	75	80
Gly	Lys	Val	Thr	Arg	Lys	Leu	Ser	Ser	Ser	Asp	Ala	Pro	Ala	Gln	Asp	85	90	95	
Thr	Gly	Ser	Ser	Ala	Ala	Ala	Val	Glu	Thr	Asp	Ala	Ser	Arg	Thr	Gly	100	105	110	
Ala	Gln	Gln	Lys	Met	Arg	Ile	Pro	Val	Ala	Arg	Val	Gln	Ala	Leu	Pro	115	120	125	
Thr	Pro	Ala	Thr	Asn	Gly	Asn	Arg	Lys	Lys	Tyr	Gln	Arg	Lys	Gly	Leu	130	135	140	
Thr	Gly	Arg	Val	Phe	Ile	Ser	Lys	Thr	Ala	Arg	Met	Lys	Trp	Gln	Leu	145	150	155	160
Leu	Glu	Arg	Arg	Val	Thr	Asp	Ile	Ile	Met	Gln	Lys	Met	Thr	Ile	Ser	165	170	175	
Asn	Met	Glu	Ala	Asp	Met	Asn	Arg	Leu	Leu	Lys	Gln	Arg	Glu	Glu	Leu	180	185	190	

Thr Lys Arg Arg Glu Lys Leu Ser Lys Arg Arg Glu Lys Ile Val Lys  
 195 200 205  
 Glu Asn Gly Glu Gly Asp Lys Asn Val Ala Asn Ile Asn Glu Glu Met  
 210 215 220  
 Glu Ser Leu Thr Ala Asn Ile Asp Tyr Ile Asn Asp Ser Ile Ser Asp  
 225 230 235 240  
 Cys Gln Ala Asn Ile Met Gln Met Glu Glu Ala Lys Glu Glu Gly Glu  
 245 250 255  
 Thr Leu Asp Val Thr Ala Val Ile Asn Ala Cys Thr Leu Thr Glu Ala  
 260 265 270  
 Arg Tyr Leu Leu Asp His Phe Leu Ser Met Gly Ile Asn Lys Gly Leu  
 275 280 285  
 Gln Ala Ala Gln Lys Glu Ala Gln Ile Lys Val Leu Glu Gly Arg Leu  
 290 295 300  
 Lys Gln Thr Glu Ile Thr Ser Ala Thr Gln Asn Gln Leu Leu Phe His  
 305 310 315 320  
 Met Leu Lys Glu Lys Ala Glu Leu Asn Pro Glu Leu Asp Ala Leu Leu  
 325 330 335  
 Gly His Ala Leu Gln Asp Leu Asp Ser Val Pro Leu Glu Asn Val Glu  
 340 345 350  
 Asp Ser Thr Asp Glu Asp Ala Pro Leu Asn Ser Pro Gly Ser Glu Gly  
 355 360 365  
 Ser Thr Leu Ser Ser Asp Leu Met Lys Leu Cys Gly Glu Val Lys Pro  
 370 375 380  
 Lys Asn Lys Ala Arg Arg Arg Thr Thr Thr Gln Met Glu Leu Leu Tyr  
 385 390 395 400  
 Ala Asp Ser Ser Glu Leu Ala Ser Asp Thr Ser Thr Gly Asp Ala Ser  
 405 410 415  
 Leu Pro Gly Pro Leu Thr Pro Val Ala Glu Gly Gln Glu Ile Gly Met  
 420 425 430  
 Asn Thr Glu Thr Ser Gly Thr Ser Ala Arg Glu Lys Glu Leu Ser Pro  
 435 440 445  
 Pro Pro Gly Leu Pro Ser Lys Ile Gly Ser Ile Ser Arg Gln Ser Ser  
 450 455 460  
 Leu Ser Glu Lys Lys Ile Pro Glu Pro Ser Pro Val Thr Arg Arg Lys  
 465 470 475 480  
 Ala Tyr Glu Lys Ala Glu Lys Ser Lys Ala Lys Glu Gln Lys His Ser

485										490					495				
Asp	Ser	Gly	Thr	Ser	Glu	Ala	Ser	Leu	Ser	Pro	Pro	Ser	Ser	Pro	Pro				
			500					505					510						
Ser	Arg	Pro	Arg	Asn	Glu	Leu	Asn	Val	Phe	Asn	Arg	Leu	Thr	Val	Ser				
		515					520					525							
Gln	Gly	Asn	Thr	Ser	Val	Gln	Gln	Asp	Lys	Ser	Asp	Glu	Ser	Asp	Ser				
	530					535					540								
Ser	Leu	Ser	Glu	Val	His	Ser	Arg	Ser	Ser	Arg	Arg	Gly	Ile	Ile	Asn				
545					550					555					560				
Pro	Phe	Pro	Ala	Ser	Lys	Gly	Ile	Arg	Ala	Phe	Pro	Leu	Gln	Cys	Ile				
				565					570					575					
His	Ile	Ala	Glu	Gly	His	Thr	Lys	Ala	Val	Leu	Cys	Val	Asp	Ser	Thr				
			580					585					590						
Asp	Asp	Leu	Leu	Phe	Thr	Gly	Ser	Lys	Asp	Arg	Thr	Cys	Lys	Val	Trp				
		595					600					605							
Asn	Leu	Val	Thr	Gly	Gln	Glu	Ile	Met	Ser	Leu	Gly	Gly	His	Pro	Asn				
	610					615					620								
Asn	Val	Val	Ser	Val	Lys	Tyr	Cys	Asn	Tyr	Thr	Ser	Leu	Val	Phe	Thr				
625					630					635					640				
Val	Ser	Thr	Ser	Tyr	Ile	Lys	Val	Trp	Asp	Ile	Arg	Asp	Ser	Ala	Lys				
				645					650					655					
Cys	Ile	Arg	Thr	Leu	Thr	Ser	Ser	Gly	Gln	Val	Thr	Leu	Gly	Asp	Ala				
			660					665					670						
Cys	Ser	Ala	Ser	Thr	Ser	Arg	Thr	Val	Ala	Ile	Pro	Ser	Gly	Glu	Asn				
		675					680					685							
Gln	Ile	Asn	Gln	Ile	Ala	Leu	Asn	Pro	Thr	Gly	Thr	Phe	Leu	Tyr	Ala				
	690					695					700								
Ala	Ser	Gly	Asn	Ala	Val	Arg	Met	Trp	Asp	Leu	Lys	Arg	Phe	Gln	Ser				
705					710					715					720				
Thr	Gly	Lys	Leu	Thr	Gly	His	Leu	Gly	Pro	Val	Met	Cys	Leu	Thr	Val				
				725					730					735					
Asp	Gln	Ile	Ser	Ser	Gly	Gln	Asp	Leu	Ile	Ile	Thr	Gly	Ser	Lys	Asp				
			740					745					750						
His	Tyr	Ile	Lys	Met	Phe	Asp	Val	Thr	Glu	Gly	Ala	Leu	Gly	Thr	Val				
		755					760						765						
Ser	Pro	Thr	His	Asn	Phe	Glu	Pro	Pro	His	Tyr	Asp	Gly	Ile	Glu	Ala				
		770				775						780							

Leu Thr Ile Gln Gly Asp Asn Leu Phe Ser Gly Ser Arg Asp Asn Gly		
785	790	795 800
Ile Lys Lys Trp Asp Leu Thr Gln Lys Asp Leu Leu Gln Gln Val Pro		
	805	810 815
Asn Ala His Lys Asp Trp Val Cys Ala Leu Gly Val Val Pro Asp His		
	820	825 830
Pro Val Leu Leu Ser Gly Cys Arg Gly Gly Ile Leu Lys Val Trp Asn		
	835	840 845
Met Asp Thr Phe Met Pro Val Gly Glu Met Lys Gly His Asp Ser Pro		
	850	855 860
Ile Asn Ala Ile Cys Val Asn Ser Thr His Ile Phe Thr Ala Ala Asp		
865	870	875 880
Asp Arg Thr Val Arg Ile Trp Lys Ala Arg Asn Leu Gln Asp Gly Gln		
	885	890 895
Ile Ser Asp Thr Gly Asp Leu Gly Glu Asp Ile Ala Ser Asn		
	900	905 910

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oignonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TNCTCGGTTAT ATGGAGGACG AATAGACT

29

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CTCTTTATGA GCATGATATG GCTTCAG

27

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNTGCTGCCCG ACTTAACTG AGAACCAA

29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GNGGACAATC CCATAGTTGA GGT TTGTC

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNGGAAGAGTG ACCCAGGGGC CCAAAGCA

29

## (2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CNCTGCAAGCG AGTGGTTTTC AGGGCTGT

29

## (2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNGAGCCAAGG AGAGAGCAGA CAGGCTGA

29

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ANCCCCCAGAA GTACAAGGAC ATGCCAAG

29



## (2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CNATAGCATCA GCATCAACAT GGTGACAA

29

## (2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 27 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CTCCATTCTC CTTGACTATC TTCTCCC

27

## (2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TNCCTGGTGCC ATCATGAT\*G TATGAGGC

29

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 22 to nucleotide 462;
  - (c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AJ1\_1 deposited under accession number ATCC 98278;
  - (d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AJ1\_1 deposited under accession number ATCC 98278;
  - (e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AJ1\_1 deposited under accession number ATCC 98278;
  - (f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AJ1\_1 deposited under accession number ATCC 98278;
  - (g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
  - (h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity;
  - (i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;
  - (j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and
  - (k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).
2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.
3. A host cell transformed with a composition of claim 2.
4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by a composition of claim 2, which process comprises:

- (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
- (b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
- (b) the amino acid sequence of SEQ ID NO:2 from amino acid 52 to amino acid 147;
- (c) fragments of the amino acid sequence of SEQ ID NO:2; and
- (d) the amino acid sequence encoded by the cDNA insert of clone

AJ1\_1 deposited under accession number ATCC 98278;  
the protein being substantially free from other mammalian proteins.

9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 52 to amino acid 147.

11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1 or SEQ ID NO:3.

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4 from nucleotide 7 to nucleotide 1647;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:4 from nucleotide 1 to nucleotide 305;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone AQ73\_3 deposited under accession number ATCC 98278;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:5;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:5 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:5;
  - (b) the amino acid sequence of SEQ ID NO:5 from amino acid 1 to amino acid 68;
  - (c) fragments of the amino acid sequence of SEQ ID NO:5; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone AQ73\_3 deposited under accession number ATCC 98278;
- the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:4.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 62 to nucleotide 757;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:6 from nucleotide 357 to nucleotide 703;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BG142\_1 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BG142\_1 deposited under accession number ATCC 98278;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:7;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:7 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:7;

(b) the amino acid sequence of SEQ ID NO:7 from amino acid 184 to amino acid 214;

(c) fragments of the amino acid sequence of SEQ ID NO:7; and

(d) the amino acid sequence encoded by the cDNA insert of clone BG142\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:6.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 404 to nucleotide 535;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:8 from nucleotide 1 to nucleotide 666;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV66\_1 deposited under accession number ATCC 98278;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV66\_1 deposited under accession number ATCC 98278;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV66\_1 deposited under accession number ATCC 98278;

- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV66\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:9;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:9 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:9;
- (b) the amino acid sequence of SEQ ID NO:9 from amino acid 1 to amino acid 38;
- (c) fragments of the amino acid sequence of SEQ ID NO:9; and
- (d) the amino acid sequence encoded by the cDNA insert of clone BV66\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:8.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 1204 to nucleotide 1389;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:10 from nucleotide 881 to nucleotide 1380;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone BV291\_3 deposited under accession number ATCC 98278;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:11;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:11 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:11;

(b) the amino acid sequence of SEQ ID NO:11 from amino acid 1 to amino acid 59;

(c) fragments of the amino acid sequence of SEQ ID NO:11; and

(d) the amino acid sequence encoded by the cDNA insert of clone BV291\_3 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:10.



26. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 189 to nucleotide 1115;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:12 from nucleotide 1 to nucleotide 451;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CK201\_1 deposited under accession number ATCC 98278;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:13;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:13 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:13;
- (b) the amino acid sequence of SEQ ID NO:13 from amino acid 1 to amino acid 88;

- (c) fragments of the amino acid sequence of SEQ ID NO:13; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone CK201\_1 deposited under accession number ATCC 98278;
- the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:12.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 117 to nucleotide 923;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 174 to nucleotide 923;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 1 to nucleotide 316;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CQ331\_2 deposited under accession number ATCC 98278;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CQ331\_2 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:15;
- (b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 57;
- (c) fragments of the amino acid sequence of SEQ ID NO:15; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CQ331\_2 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:14.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 223 to nucleotide 483;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 22 to nucleotide 397;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT550\_1 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT550\_1 deposited under accession number ATCC 98278;

- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:17;
- (b) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 58;
- (c) fragments of the amino acid sequence of SEQ ID NO:17; and
- (d) the amino acid sequence encoded by the cDNA insert of clone CT550\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:16.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 112 to nucleotide 969;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 154 to nucleotide 969;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 1 to nucleotide 423;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT585\_1 deposited under accession number ATCC 98278;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 1 to amino acid 104;

(c) fragments of the amino acid sequence of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone CT585\_1 deposited under accession number ATCC 98278;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:18.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 37 to nucleotide 2766;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 243 to nucleotide 789;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CT797\_3 deposited under accession number ATCC 98278;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CT797\_3 deposited under accession number ATCC 98278;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CT797\_3 deposited under accession number ATCC 98278;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

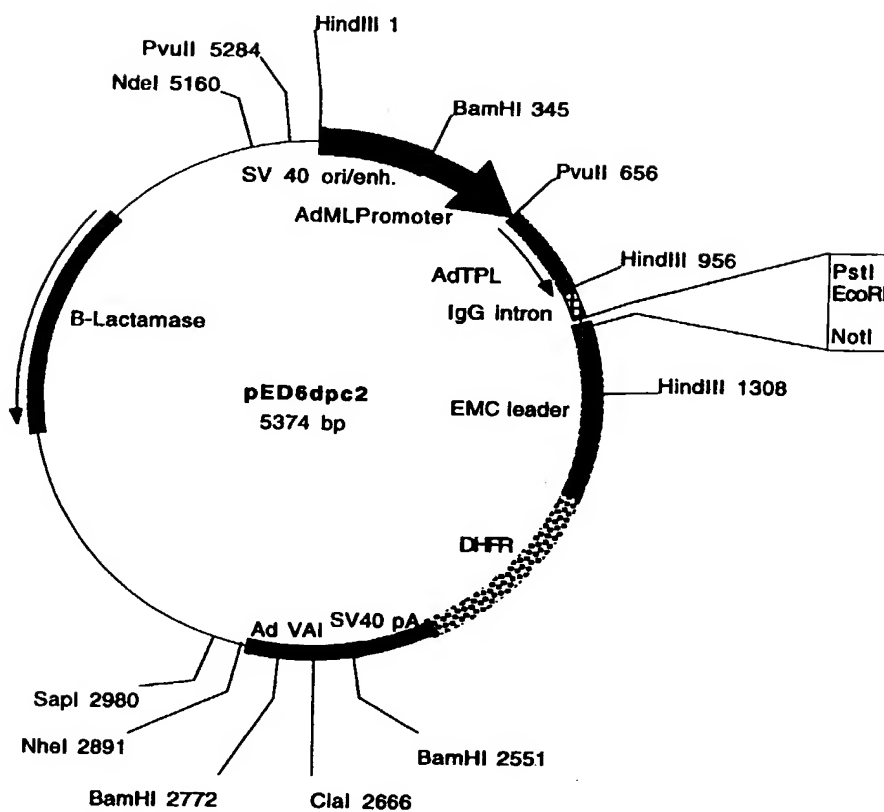
39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:21;
- (b) the amino acid sequence of SEQ ID NO:21 from amino acid 75 to amino acid 251;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and  
(d) the amino acid sequence encoded by the cDNA insert of clone  
CT797\_3 deposited under accession number ATCC 98278;  
the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

FIGURE 1A



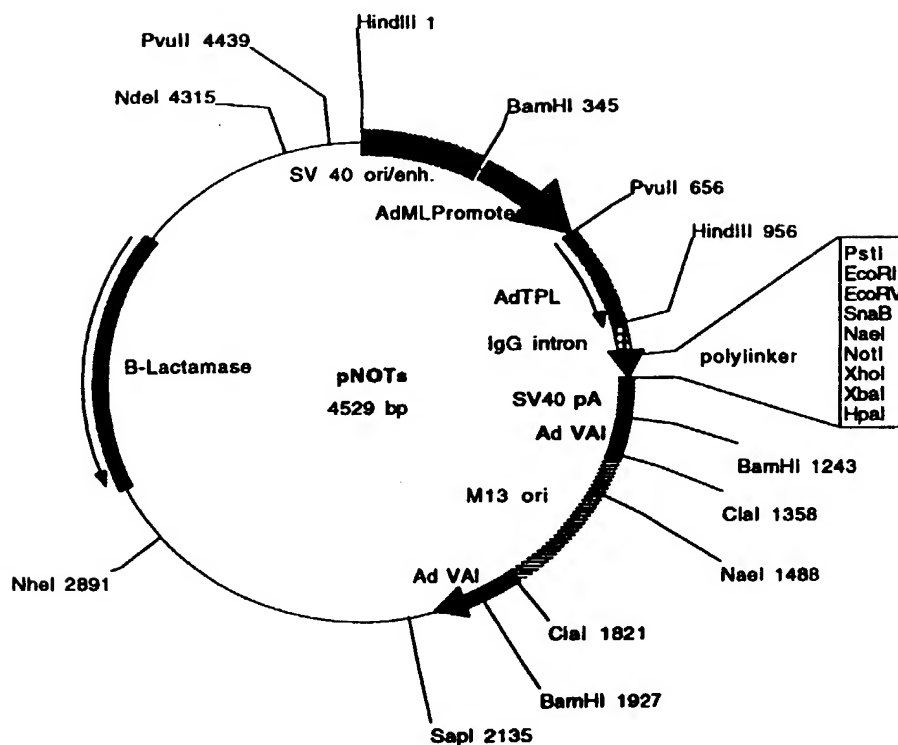
**Plasmid name:** pED6dpc2

**Plasmid size:** 5374 bp

**Comments/References:** pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.



FIGURE 1B



**Plasmid name:** pNOTs

**Plasmid size:** 4529 bp

**Comments/References:** pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI